Title

Characterizing the relationship between steady state and response using analytical expressions for the steady states of mass action models.

Journal

PLoS computational biology, 9(2)

ISSN

1553-734X

Authors

Loriaux, Paul Michael
Tesler, Glenn
Hoffmann, Alexander

Publication Date

2013

DOI

10.1371/journal.pcbi.1002901

Peer reviewed

OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

Characterizing the Relationship between Steady State and Response Using Analytical Expressions for the Steady States of Mass Action Models

Paul Michael Loriaux[1,2,3], Glenn Tesler[4], Alexander Hoffmann[1,3]*

1 Signaling Systems Laboratory, Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, California, United States of America, **2** Graduate Program in Bioinformatics and Systems Biology, University of California San Diego, La Jolla, California, United States of America, **3** The San Diego Center for Systems Biology, La Jolla, California, United States of America, **4** Department of Mathematics, University of California San Diego, La Jolla, California, United States of America

Abstract

The steady states of cells affect their response to perturbation. Indeed, diagnostic markers for predicting the response to therapeutic perturbation are often based on steady state measurements. In spite of this, no method exists to systematically characterize the relationship between steady state and response. Mathematical models are established tools for studying cellular responses, but characterizing their relationship to the steady state requires that it have a parametric, or analytical, expression. For some models, this expression can be derived by the King-Altman method. However, King-Altman requires that no substrate act as an enzyme, and is therefore not applicable to most models of signal transduction. For this reason we developed *py*-substitution, a simple but general method for deriving analytical expressions for the steady states of mass action models. Where the King-Altman method is applicable, we show that *py*-substitution yields an equivalent expression, and at comparable efficiency. We use *py*-substitution to study the relationship between steady state and sensitivity to the anti-cancer drug candidate, dulanermin (recombinant human TRAIL). First, we use *py*-substitution to derive an analytical expression for the steady state of a published model of TRAIL-induced apoptosis. Next, we show that the amount of TRAIL required for cell death is sensitive to the steady state concentrations of procaspase 8 and its negative regulator, Bar, but not the other procaspase molecules. This suggests that activation of caspase 8 is a critical point in the death decision process. Finally, we show that changes in the threshold at which TRAIL results in cell death is not always equivalent to changes in the time of death, as is commonly assumed. Our work demonstrates that an analytical expression is a powerful tool for identifying steady state determinants of the cellular response to perturbation. All code is available at http://signalingsystems.ucsd.edu/models-and-code/ or as supplementary material accompanying this paper.

Citation: Loriaux PM, Tesler G, Hoffmann A (2013) Characterizing the Relationship between Steady State and Response Using Analytical Expressions for the Steady States of Mass Action Models. PLoS Comput Biol 9(2): e1002901. doi:10.1371/journal.pcbi.1002901

Editor: Daniel A. Beard, Medical College of Wisconsin, United States of America

Received August 17, 2012; **Accepted** December 12, 2012; **Published** February 28, 2013

Funding: The study was supported by the NIH via R01 grant GM072024 and P50 GM085763, the San Diego Center for Systems Biology. PML was supported by the Department of Energy Computational Sciences Graduate Fellowship. GT was supported by the National Science Foundation (NSF grants DMS-0718810 and CCF-1115206). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: ahoffmann@ucsd.edu

Introduction

Transient activation of signaling molecules is a hallmark of the cellular response to perturbation. Far from acting as a simple relay, however, the dynamics of signaling molecules can encode information about the instigating stimulus [1–3]. Interestingly, these dynamics are affected by the steady state prior to perturbation [4,5]. Non-genetic variation in the proteome, for example, is sufficient to explain variability in the sensitivity of HeLa cells to the pro-apoptotic ligand TRAIL [6]. Like other TNF superfamily members, TRAIL is a promising anti-cancer therapeutic [7]. Recombinant human TRAIL, or dulanermin, as well as antibodies raised against the TRAIL receptors DR4 and DR5, are currently in clinical trials [8]. To improve the efficacy of these and other drugs, understanding how sensitivity is affected by the cellular resting state is of great importance [9].

Mathematical models are powerful tools for characterizing the behavior of signaling systems in response to perturbation [10–13]. Assuming conservation of mass, these models equate the change in concentration of a molecular species with the sum of reaction velocities that produce the species, minus the sum of those that consume it. The reactions themselves are often modeled by the *Law of Mass Action*. This law assumes that the velocity of a reaction is proportional to the product of the concentrations of its reactants. Since many signaling reactions are bimolecular, the resulting *mass balance* equations are non-linear in the concentrations. A system is at *steady state* if no species is consumed faster than it is produced, nor produced faster than it is consumed. By this formalism, the steady state of a signaling system is equivalent to the root of a non-linear system of equations. Because of this, no universal method has been developed to identify the steady states of mass action models, despite their importance to basic and clinical research. As a result, even with the help of mathematical models, investigating the relationship between steady state and stimulus-responsiveness remains cumbersome.

Of course with any model, simulating the response to perturbation often requires the system to be at steady state prior to perturbation. To achieve this, one of several techniques is currently used. The most common technique is to assume a

Author Summary

Diagnostic markers are derived from steady state measurements, but are used to predict the cellular response to therapy. To develop new and better diagnostics, we would like to systematically characterize the relationship between steady state and the response to a given therapeutic. Mathematical models have powerfully complemented empirical studies in this regard, but it remains challenging to employ these models to characterize the effects of steady state. To do so requires a mathematical expression for the steady state, for which no universal method has been developed. Here, we present a method for deriving a mathematical expression for the steady state of a common class of models, those that obey the Law of Mass Action. We show that our method is easy to use and scales well to large models. We then use our method to characterize the relationship between steady state and the sensitivity to the anti-cancer drug candidate, dulanermin. We find that sensitivity to the drug is strongly affected by the concentration of the signaling molecule, procaspase 8, and its inhibitor, Bar. Our work thus demonstrates the utility of analytical studies of the steady state and its relationship to drug sensitivity.

"trivial" steady state where every reaction velocity is zero [2,14]. While straightforward, this approach may not reflect biological reality, where tonic signaling is common [15,16] and can strongly influence the response to perturbation [17–19]. A second technique is to approach the steady state asymptotically via numerical integration of the mass balance equations [1,13,20]. While this approach can yield non-trivial steady states, the number of integration steps required to reach the steady state may dominate the number of steps required to simulate the perturbation. Also, identifying the parameter values that result in a desired steady state is an inverse problem that requires non-linear optimization. For these reasons, numerical derivation of the steady state is impractical when characterizing its effect on the response to perturbation, and an analytical expression is required instead.

The best-known method for deriving analytical expressions for the steady states of mass action models was developed by King and Altman in 1956 [21]. This method assumes that all molecular species can be divided into enzymes and substrates, that no enzyme is itself a substrate, and that all substrates remain constant over the time-scale of steady state formation [22]. A number of improvements have been made to the King-Altman method over the years [23–25]. Many of these are now implemented in the Matlab application, *KAPattern* [26]. The King-Altman methodology was also recently formalized using concepts from algebraic geometry [22,27], and extended to layered signaling cascades [28] and post-translational modification networks [29]. Despite these improvements, however, these methods do not extend to mass action models with arbitrary reaction structure, as is common in contemporary models of signaling systems. Furthermore, only the King-Altman method has been reduced to practice.

For these reasons we developed *py*-substitution, a simple, algebraic method for deriving steady state expressions for mass action models with arbitrary structure. Our method can be explained using concepts from linear algebra, and full code has been provided for all examples in this manuscript, implemented in either Matlab or Maple. A particular benefit of *py*-substitution is that it affords considerable flexibility when selecting independent quantities for the steady state expression. Often, this permits explicit derivation of kinetic rate constants from steady state concentration measurements. More generally, it allows independent quantities to be chosen that maximize incorporation of known or measured parameter values. This not only simplifies model fitting, but typically reduces the total number of parameters required as well. We compare *py*-substitution to the King-Altman method and show that, where King-Altman is applicable, the two methods yield equivalent results. Computationally, however, we find that our method is more efficient, and, because *py*-substitution does not require a particular reaction structure, more general than King-Altman.

Finally, we use *py*-substitution to derive a steady state expression for a recent model of apoptosis induced by the death-receptor ligand TRAIL [14]. We find that incorporation of a non-trivial steady state changes the qualitative behavior of the model. Specifically, tonic signaling desensitizes the system to low doses of TRAIL, while high doses of TRAIL still result in the "snap-action" signaling dynamics indicative of cell death. We then systematically alter the steady state and show that changes in steady state affect the threshold at which TRAIL results in death. We find that the threshold is highly sensitive to the steady state abundances of procaspase 8 and its negative regulator, Bar, but not the other procaspase molecules. This suggests that the activation of caspase 8 is a critical point in the cell death decision. Finally, without recourse to a model that is tolerant to low doses of TRAIL, a common practice is to approximate the sensitivity to TRAIL by the time at which death occurs. Using our tonic signaling model, we show that these two metrics are not universally equivalent. Caution should therefore be taken when equating the dynamics of cell death with the probability that death occurs.

Materials and Methods

In this section we describe the process for deriving analytical expressions for the steady states of mass action models using *py*-substitution. First we describe the class of models to which *py*-substitution can be applied. Next, we review existing methods for deriving analytical expressions for the steady states of these models. Finally, we describe *py*-substitution using some formal concepts from algebra. In the results section we provide several examples, beginning with a version of the classical Michaelis-Menten model of enzyme action. All code for these examples, as well as detailed instructions for use and full transcripts of the output, are provided in Protocol S1 and on our website, http://signalingsystems.ucsd.edu/models-and-code/.

Preliminaries

Let $\mathbb{N}_0$ be the set of non-negative natural numbers and $\mathbb{R}_0$ be the set of non-negative real numbers. Let $\mathcal{A}=\{a_1,a_2,\ldots,a_{d_x}\}$ be a set of d_x species and $\mathcal{R}=\{r_1,r_2,\ldots,r_{d_k}\}$ be a set of d_k reactions. Each reaction $r_j\in\mathcal{R}$ follows the normal definition,

$$r_j: s^{in}_{1,j}a_1+s^{in}_{2,j}a_2+\cdots+s^{in}_{d_x,j}a_{d_x} \xrightarrow{v_j} s^{out}_{1,j}a_1+s^{out}_{2,j}a_2+\cdots s^{out}_{d_x,j}a_{d_x}$$

where $s^{in}_{i,j}\in\mathbb{N}_0$ is the stoichiometric coefficient of the i^{th} reactant and $s^{out}_{i,j}\in\mathbb{N}_0$ is the stoichiometric coefficient of the i^{th} product [30]. We define $x_i\in\mathbb{R}_0$ to be the concentration of species a_i and $v_j\in\mathbb{R}_0$ to be the velocity at which r_j converts reactants into products. By the *Law of Mass Action*,

$$v_j = k_j \prod_{i=1}^{d_x} x_i^{h_{i,j}}. \quad (1)$$

The quantity $h_{i,j} \in \mathbb{R}_0$ is often, but not necessarily, equal to $s_{i,j}^{in}$. The coefficient $k_j \in \mathbb{R}_0$ is called the *rate constant*. Assuming conservation of mass, the concentration x_i changes according to

$$\dot{x}_i = \sum_{j=1}^{d_k} (s_{i,j}^{out} - s_{i,j}^{in}) v_j, \quad (2)$$

where $\dot{x}_i$ is the first derivative of x_i with respect to time. Any collection $\{\mathcal{A},\mathcal{R}\}$ where the concentration x_i of $a_i \in \mathcal{A}$ obeys Equation 2 and the velocity v_j of $r_j \in \mathcal{R}$ obeys Equation 1 is called a *mass action model*. In what follows, we assume i, i_1, and i_2 are indices over the interval $1,\ldots,d_x$ and j is an index over $1,\ldots,d_k$. When $v_1,\ldots,v_{d_k}$ are such that all

$$\dot{x}_i = 0, \quad (3)$$

the model is said to be at *steady state*. If all $v_j = 0$ we call the steady state *trivial*. In this manuscript we are concerned with symbolic, non-trivial solutions to Equation 3. A solution is symbolic if all k_j and x_i are left as uninterpreted variables, rather than being assigned numerical values. For a complete list of symbols and their meanings, see Table S1.

Prior work

Let $\mathbf{x} \in \mathbb{R}_0^{d_x}$ and $\mathbf{v} \in \mathbb{R}_0^{d_k}$ be the vectors with elements $(\mathbf{x})_i = x_i$, and $(\mathbf{v})_j = v_j$. Throughout this manuscript, we use $(\mathbf{b})_i$ to denote the i^{th} element of vector $\mathbf{b}$ and $(\mathbf{A})_{ij}$ to denote the element at row i, column j of matrix $\mathbf{A}$. Let $\mathbf{S}$ be the *stoichiometric matrix*, i.e., the matrix whose elements are $(\mathbf{S})_{ij} = s_{i,j}^{out} - s_{i,j}^{in}$. Using this notation, Equation 2 becomes

$$\dot{\mathbf{x}} = \mathbf{S}\mathbf{v}, \quad (4)$$

and the steady state equation becomes

$$\mathbf{S}\bar{\mathbf{v}} = 0. \quad (5)$$

By convention we use the overline to denote vectors that satisfy steady state. Equation 5 often takes this form in flux balance analysis [31–34]. Here $\bar{\mathbf{v}}$ is a real-valued vector and is calculated numerically. However, prior work has shown that Equation 5 can also be used to calculate a vector of rate constants from a vector of steady state concentrations [35]. Let $\mathbf{k} \in \mathbb{R}_0^{d_k}$ be the vector with elements $(\mathbf{k})_j = k_j$. Let $\mathbf{P}_k$ be the diagonal matrix with elements $(\mathbf{P}_k)_{j,j} = (\partial v_j / \partial k_j)$. The vector $\mathbf{v}$ can then be expressed as

$$\mathbf{v} = \mathbf{P}_k \mathbf{k}. \quad (6)$$

Substituting Equation 6 into Equation 5 and solving for $\mathbf{k}$ yields the k-cone [35] — equivalently, the left null space of the matrix product $\mathbf{S}\mathbf{P}_k$. Given a basis for this null space and a vector of steady state concentrations, a vector of rate constants can be calculated that satisfies Equation 5. While this approach is useful for deriving kinetic parameters from metabolomic measurements, it is less well suited to signaling systems where transient and low-abundance species confound accurate measurement of the concentrations.

If the velocity of every $r_j \in \mathcal{R}$ is homogeneous of degree 1 in $x_1,\ldots,x_{d_x}$, then an analogous approach allows $\mathbf{v}$ to be expressed in terms of $\mathbf{x}$. We call models that satisfy this condition *linear models*. An alternative, stoichiometric definition for a linear model is given by the following,

$$\forall r_j \in \mathcal{R},\ \sum_{i_1=1}^{d_k} s_{i_1,j}^{in} = \sum_{i_2=1}^{d_k} s_{i_2,j}^{out} = 1. \quad (7)$$

Equation 7 requires that every reaction defines a transition from exactly one time-varying species to another. Let $\mathbf{P}_x$ be the matrix with elements $(\mathbf{P})_{i,j} = (\partial v_i / \partial x_j)$. If $\mathbf{v}$ is a vector of linear reaction velocities, it can likewise be expressed as

$$\mathbf{v} = \mathbf{P}_x \mathbf{x}. \quad (8)$$

Substituting Equation 8 into Equation 5 results in the matrix product $\mathbf{S}\mathbf{P}_x$, also called the *Jacobian matrix* [36]. Given a basis for the null space of the Jacobian, a vector of steady state concentrations can be calculated from a vector rate constants.

For linear models, an alternative, graphical method for deriving expressions for the steady state species concentrations was introduced by King and Altman in 1956 [21]. Notice that Equation 7 permits a two-dimensional indexing of the rate constants,

$$k'_{i_1,i_2} = \begin{cases} k_j & if\ \exists\ r_j \in \mathcal{R} : s_{i_1,j}^{in} = s_{i_2,j}^{out} = 1 \\ 0 & otherwise. \end{cases} \quad (9)$$

We call k'_{i_1,i_2} a *transition* rate constant since the product $k'_{i_1,i_2} x_{i_1}$ defines the rate of transition from species x_{i_1} to x_{i_2}. Substituting Equation 9 into Equations 1 and 2 gives

$$\frac{dx_{i_1}}{dt} = \sum_{i_2=1}^{d_x} k'_{i_2,i_1} x_{i_2} - x_{i_1} \sum_{i_2=1}^{d_x} k'_{i_1,i_2}. \quad (10)$$

By defining the matrix $\mathbf{K}$ with elements

$$(\mathbf{K})_{i_1,i_2} = \begin{cases} k'_{i_2,i_1} & if\ i_1 \neq i_2, \\ \sum_{m \neq i_1} -k'_{i_1,m} & if\ i_1 = i_2, \end{cases} \quad (11)$$

the steady state equation becomes

$$\mathbf{K}\mathbf{x} = 0. \quad (12)$$

Note that $\mathbf{K}$ is simply the Jacobian matrix for a linear model, $\mathbf{K} = \mathbf{S}\mathbf{P}_x$. The general solution to Equation 12 was found in [21] to be the vector $\bar{\mathbf{x}}$ with elements

$$(\bar{\mathbf{x}})_i = \frac{\mathbf{M}_i}{\sum_{i_2=1}^{d_x} \mathbf{M}_{i_2}}, \tag{13}$$

where $\mathbf{M}_i$ is the i^{th} minor of $\mathbf{K}$, formed by removing its i^{th} row and column and computing its determinant. For sufficiently small systems, Equation 13 can be solved directly using modern mathematical computing software [37]. Prior to the advent of modern computers, King and Altman realized that the minors can also be derived by graph theoretic means. Note that for a linear model, $\mathcal{A}$ and $\mathcal{R}$ imply a directed graph,

$$\mathcal{G} = (\mathcal{A}, \mathcal{R}), \tag{14}$$

where each $a_i \in \mathcal{A}$ defines a vertex and each $r_j \in \mathcal{R}$ defines an edge between vertices a_{i_1} and a_{i_2} (provided i_1 and i_2 are such that $s^{in}_{i_1,j} = s^{out}_{i_2,j} = 1$). The King-Altman method enumerates for each species $a_i \in \mathcal{A}$ the set $\mathcal{S}_i$ of simple connected subgraphs

$$\mathcal{S}_i = \{\mathcal{G}' = (\mathcal{X}, \mathcal{R}') : \mathcal{R}' \subset \mathcal{R}, |\mathcal{R}'| = d_x - 1\}$$

where vertex a_i has out-degree 0 and all other vertices have out-degree 1 [23,24]. These are the directed spanning trees of G, with all edges directed towards root a_i. A subgraph $\mathcal{G}'$ is called a King-Altman *pattern*. The minor $\mathbf{M}_i$ can then be expressed as

$$\mathbf{M}_i = \sum_{\mathcal{G}' \in \mathcal{S}_i} \prod_{r_j \in \mathcal{R}'} k_j, \tag{15}$$

where $k_j = k'_{i_1,i_2}$ is the transition rate constant between species a_{i_1} and a_{i_2}. For a more thorough derivation of the King-Altman method, see [38].

Of course, many biochemical reactions are bimolecular. By Equation 1, the velocity of a bimolecular reaction is degree 2 in $x_1, \ldots, x_{d_x}$. To preserve linearity, one can assume the concentration of one reactant is so high as to be effectively constant. This concentration is incorporated into the kinetic rate constant, and the techniques described above can still be used to solve Equation 3. If this assumption fails, then Equation 2 describes a polynomial in $x_1, \ldots, x_{d_x}$ with coefficients in $\mathbb{Q}[k_1, \ldots, k_{d_k}]$. In this case the solutions to Equation 3 form an algebraic variety. Deriving an expression for the steady state of a non-linear model thus requires finding a parameterization of the variety [39]. One way to achieve this is to calculate a Gröbner basis for the ideal generated by $\dot{\boldsymbol{x}}_1, \ldots, \dot{\boldsymbol{x}}_{d_k}$ and eliminate variables [40,41]. Alternatively, if the model displays certain structural properties, variables can be eliminated by identifying conservation relationships. The best-known example of this is when $\{\mathcal{A}, \mathcal{R}\}$ defines a cascade of post-translational modifications. In this case, enzyme-substrate intermediates can be eliminated and the variety can be parameterized by rational functions of the free enzyme concentrations with coefficients in $\mathbb{Q}(k_1, \ldots, k_{d_k})$ [22,28]. Although these methods do not require linearity, calculating a Gröbner basis can be computationally intractable, while identifying conservation relationships can be difficult for models of arbitrary reaction structure.

py-substitution

Py-substitution allows mass action models — a particular class of non-linear model — to be solved using simple linear algebra. We make use of the following observations: (a) $\dot{\boldsymbol{x}}_i$ is always homogeneous of degree 1 in $k_1, \ldots, k_{d_k}$, and (b) $\dot{\boldsymbol{x}}_i$ is often no greater than degree 2 in $x_1, \ldots, x_{d_x}$. If a subset of elements in $\mathcal{K} \cup \mathcal{X}$ can be found on which every $\dot{\boldsymbol{x}}_i$ has only linear dependence, then Equation 5 can be solved using linear methods.

To begin, we define sets of symbolic variables $\mathcal{P} = \{p_1, \ldots, p_{d_p}\}$ and $\mathcal{Y} = \{y_1, \ldots, y_{d_y}\}$ such that $d_p + d_y = d_k + d_x$ and $d_y \geq \text{rank } \mathbf{S}$. We then relabel, or map, every element in $\mathcal{K} \cup \mathcal{X}$ to a unique element in $\mathcal{P} \cup \mathcal{Y}$ so that every $\dot{\boldsymbol{x}}_i$ is linear in $\mathcal{Y}$. By Equations 1 and 2 this requires that all v_j are linear in $\mathcal{Y}$. Variables that we want to remain independent, as well as variables on which $\dot{\boldsymbol{x}}_i$ has non-linear dependence, should be mapped to $\mathcal{P}$. As we shall see, there is considerable flexibility in choosing this map.

Let $\mathcal{K}$ and $\mathcal{X}$ be partitioned into disjoint (but possibly empty) subsets $\mathcal{K} = \mathcal{K}_p \cup \mathcal{K}_{lin}$ and $\mathcal{X} = \mathcal{X}_p \cup \mathcal{X}_{lin}$. We define ψ_p to be a bijective map (with a restriction given below)

$$\psi_p : \begin{cases} \mathcal{K}_p \cup \mathcal{X}_p \rightarrow \mathcal{P} \\ \mathcal{K}_{lin} \cup \mathcal{X}_{lin} \rightarrow \mathcal{Y}, \end{cases}$$

and extend it homomorphically over $\mathbb{Q}[\mathcal{K}][\mathcal{X}]$. Our linearity restriction is to consider maps of this form such that

$$\psi_p(v_j) = y_n \prod_{m=1}^{d_p} p_m^{h'_{j,m}} \tag{16}$$

for some $y_n \in \mathcal{Y}$. For $p_m = \psi_p(x_i)$, the exponent is $h'_{j,m} = h_{i,j}$. For $p_m = \psi_p(k_j)$, the exponent $h'_{j,m} = 1$. In words, ψ_p defines a change of variables such that $\psi_p(v_j)$ is homogeneous of degree 1 in $y_1, \ldots, y_{d_y}$. By Equation 2, $\psi_p(\dot{\boldsymbol{x}}_i)$ becomes a homogeneous polynomial of degree 1 in $y_1, \ldots, y_{d_y}$ with coefficients in $\mathbb{Q}[p_1, \ldots, p_{d_p}]$. We can now write

$$\psi_p(\mathbf{v}) = \mathbf{P}\mathbf{y}, \tag{17}$$

where $\mathbf{P}$ is the $d_k \times d_y$ Jacobian matrix with elements $(\mathbf{P})_{ij} = (\partial v_i / \partial y_j)$. Here and elsewhere we use the notation $\psi(\mathbf{v}) = \mathbf{w}$ to mean that $\mathbf{w}$ is the vector formed by applying the function ψ element-wise to $\mathbf{v}$. Note that the trivial partition $\mathcal{K}_{lin} = \mathcal{K}$ and $\mathcal{X}_p = \mathcal{X}$ recovers the k-cone procedure described above. For the remainder of this section, we treat j as an index over $1, \ldots, d_y$. Substituting Equation 17 into Equation 5 gives

$$\mathbf{C}\mathbf{y} = 0 \tag{18}$$

where $\mathbf{C} = \mathbf{S}\mathbf{P}$ is called the *coefficient matrix*. The solution to Equation 18 is precisely the null space of $\mathbf{C}$. Let $\mathbf{N}$ be a matrix whose columns form a basis for this null space. Let d_q be the number of columns in $\mathbf{N}$. By the rank-nullity theorem, we have

$$d_q = \text{ncols } \mathbf{C} - \text{rank } \mathbf{C}, \tag{19}$$

where $ncols\ \mathbf{C} = d_y$ is the number of columns in $\mathbf{C}$. Furthermore, because $\psi_p(\mathbf{v})$ is linear in $\mathbf{y}$ and ψ_p^{-1} exists, the matrix $\mathbf{P}$ must be full rank. By the properties of the rank, we can write

$$\text{rank } \mathbf{C} = \text{rank } \mathbf{SP} = \text{rank } \mathbf{S}. \tag{20}$$

Together, Equations 19 and 20 give

$$d_q = d_y - \text{rank } \mathbf{S}, \tag{21}$$

thus calling for the constraint $d_y \geq \text{rank}\mathbf{S}$. This, in conjunction with Equation 16, are the only constraints on ψ_p. If we now let $\bar{\mathbf{y}}$ be some linear combination of the basis vectors,

$$\bar{\mathbf{y}} = \mathbf{Nq}, \tag{22}$$

then $\bar{\mathbf{y}}$ satisfies Equation 18 and steady state is achieved. In general, Equation 22 is underdetermined. Equation 22 therefore implies a partition of $\mathcal{Y}$ into independent variables (denoted $\mathcal{Y}_q$) and dependent variables (denoted $\mathcal{Y}_q^c$). We will now describe this partition by a second mapping function, ψ_y.

Recall that a basis for the null space of $\mathbf{C}$ can be constructed from $\mathbf{C}_{\text{rref}}$, the reduced row echelon form of $\mathbf{C}$. Let $\mathbf{c}_j$ be the j^{th} column of $\mathbf{C}_{\text{rref}}$. If $\mathbf{c}_j$ contains a pivot position, then y_j is a dependent variable. If $\mathbf{c}_j$ does not contain a pivot, then y_j is free, or independent. Let

$$\begin{aligned} \mathcal{Y}_q &= \{y_j \in \mathcal{Y} : \text{column } \mathbf{c}_j \text{ does } \textit{not} \text{ contain a pivot}\} \\ \mathcal{Y}_q^c &= \{y_j \in \mathcal{Y} : \text{ column } \mathbf{c}_j \text{ contains a pivot}\}. \end{aligned} \tag{23}$$

Let d_q be the cardinality of $\mathcal{Y}_q$. Enumerate these variables as $\mathcal{Y}_q = \{y_{j_1}, y_{j_2}, \ldots, y_{j_{d_q}}\}$, with $j_1 < \ldots < j_{d_q}$. For every $\mathbf{c}_j$ not containing a pivot, there is a basis vector $\mathbf{n}_k$ (related by $j = j_k$) whose j^{th} element equals 1 and whose elements in positions $> j$ are 0. By Equation 22, this gives an independent parameter, $y_j = (\mathbf{q})_k = q_k$. Equation 22 thus defines a function $\psi_y : y_j \mapsto q_k$. Let $\mathcal{Q} = \{q_1, q_2, \ldots, q_{d_q}\}$ be the set of independent parameters. If column $\mathbf{c}_j$ does contain a pivot, then y_j depends on variables in $\mathcal{P} \cup \mathcal{Q}$, giving $\psi_y : y_j \mapsto f_j(\mathcal{P}, \mathcal{Q}) \in \text{span}_{\mathbb{Q}(\mathcal{P})}(\mathcal{Q})$ where $f_j(\mathcal{P}, \mathcal{Q})$ is the specific function resulting from the row operations used to reduce $\mathbf{C}$ to $\mathbf{C}_{\text{rref}}$. Equation 22 can now be described in its entirety by the mapping function ψ_y,

$$\psi_y : \begin{cases} \mathcal{P} \to \mathcal{P} \text{ (identity)} \\ \mathcal{Y}_q \to \mathcal{Q} \\ \mathcal{Y}_q^c \to \text{span}_{\mathbb{Q}(\mathcal{P})}(\mathcal{Q}). \end{cases} \tag{24}$$

The notation $\psi : \mathcal{P} \to \mathcal{P}(\text{identity})$ indicates that $\psi(p) = p$ for every $p \in \mathcal{P}$. Note that we define $\text{span}_F(\mathcal{Q})$ as the set of all linear combinations $a_1 q_1 + a_2 q_2 + \cdots$, where $a_1, a_2, \ldots \in F$ and $q_1, q_2, \ldots$ are distinct elements of $\mathcal{Q}$. $\mathbb{Q}[\mathcal{P}]$ is the set of all polynomials in variables $\mathcal{P}$ with rational numbers as coefficients. $\mathbb{Q}(\mathcal{P})$ is the field of fractions of $\mathbb{Q}[\mathcal{P}]$: any $f \in \mathbb{Q}[\mathcal{P}]$ can be expressed as $f = g_1/g_2$, where $g_1, g_2 \in \mathbb{Q}[\mathcal{P}]$.

As with ψ_p, there is some flexibility in choosing how $\mathcal{Y}$ is partitioned into free variables, $\mathcal{Y}_q$, and dependent variables, $\mathcal{Y}_q^c$. A different indexing of the variables in $\mathcal{Y}$ simultaneously permutes the vector $\mathbf{y}$ and the columns of $\mathbf{C}$. This leads to different reduced row echelon forms, with different partitions into free and dependent variables. The null space basis obtained by reducing $\mathbf{C}$ to $\mathbf{C}_{\text{rref}}$ greedily classifies low-numbered columns as dependent columns when possible, or free columns when not possible. Quantities in $\mathcal{Y}$ for which good numerical estimates exist should therefore be assigned to higher indices. These quantities are favored, but not guaranteed, to be mapped to independent parameters. Quantities for which good numerical estimates do not exist should be assigned to low indices in $\mathcal{Y}$.

Finer control over the partition of $\mathcal{Y}$ into dependent and independent parameters is possible by working directly with $\mathbf{C}_{\text{rref}}$ or $\mathbf{N}$. Let $\mathcal{Y}_q' = \{y_{j_1}, \ldots, y_{j_{d_q}}\}$ be the set of d_q elements in $\mathcal{Y}$ that we want mapped to $\mathcal{Q}$. Let $\mathbf{N}'$ be the square matrix formed by rows $j_1, \ldots, j_{d_q}$ of $\mathbf{N}$. To map $\mathcal{Y}_q'$ to $\mathcal{Q}$ requires that we find a vector $\mathbf{q}'$ such that

$$\mathbf{N}'\mathbf{q}' = \mathbf{q},$$

where $\mathbf{q}$ is the vector with elements $(\mathbf{q})_k = q_k$. Solving for $\mathbf{q}'$ gives

$$\mathbf{q}' = (\mathbf{N}')^{-1}\mathbf{q}. \tag{25}$$

Thus, for a given map ψ_p, not all partitions of $\mathcal{Y}$ into $\mathcal{Y}_q$ and $\mathcal{Y}_q^c$ are possible, but only those for which $\det(\mathbf{N}') \neq 0$. An example of this can be seen in the file "fum2.m" in Supporting Protocol S1, discussed below.

Next let $\mathcal{K}_q = \{k \in \mathcal{K} : \psi_{py}(k) \in \mathcal{Q}\}$, and $\mathcal{K}_y = \mathcal{K}_{lin} \backslash \mathcal{K}_q$. Let $\mathcal{X}_q$ and $\mathcal{X}_y$ be defined analogously. The composition $\psi_{py} = (\psi_y \circ \psi_p)$ captures the entire process of linearizing $\mathbf{v}$ with the function ψ_p, solving the linear system $\mathbf{S}\psi_p(\mathbf{v}) = 0$, and taking an arbitrary combination of solution space basis vectors:

$$\psi_{py} : \begin{cases} \mathcal{K}_p \cup \mathcal{X}_p \to \mathcal{P} \\ \mathcal{K}_q \cup \mathcal{X}_q \to \mathcal{Q} \\ \mathcal{K}_y \cup \mathcal{X}_y \to \text{span}_{\mathbb{Q}(\mathcal{P})}(\mathcal{Q}). \end{cases}$$

Applying ψ_{py} to the sets $\mathcal{K}$ and $\mathcal{X}$ results in a parametric description of the steady state that is typically the most useful: every element in $\mathcal{K}$ or $\mathcal{X}$ is mapped to an element in $\mathcal{P}$ or $\mathcal{Q}$, or a function in $\text{span}_{\mathbb{Q}(\mathcal{P})}(\mathcal{Q})$. Assigning numerical values to elements in $\mathcal{P}$ and $\mathcal{Q}$ results in elements in $\text{span}_{\mathbb{Q}(\mathcal{P})}(\mathcal{Q})$ taking values that satisfy the steady state equation. In some cases we may wish to reverse the substitution so that functions of variables $\mathcal{P} \cup \mathcal{Q}$ are mapped back to functions of $\mathcal{K} \cup \mathcal{X}$. To do so, let $\mathcal{K}_{pq} = \mathcal{K}_p \cup \mathcal{K}_q$ and $\mathcal{X}_{pq} = \mathcal{X}_p \cup \mathcal{X}_q$. Let ψ_q^{-1} be the inverse of ψ_y restricted to the independent parameters, $\mathcal{P} \cup \mathcal{Q}$.

$$\psi_q^{-1} : \begin{cases} \mathcal{P} \to \mathcal{P} \text{ (identity)} \\ \mathcal{Q} \to \mathcal{Y}_q. \end{cases}$$

The composition of ψ_p^{-1} and ψ_q^{-1} now defines a map from the set of independent parameters to their counterparts in $\mathcal{K}$ and $\mathcal{X}$,

$$\psi_{pq}^{-1} = (\psi_p^{-1} \circ \psi_q^{-1}) : \begin{cases} \mathcal{P} \to \mathcal{K}_p \cup \mathcal{X}_p \\ \mathcal{Q} \to \mathcal{K}_q \cup \mathcal{X}_q. \end{cases}$$

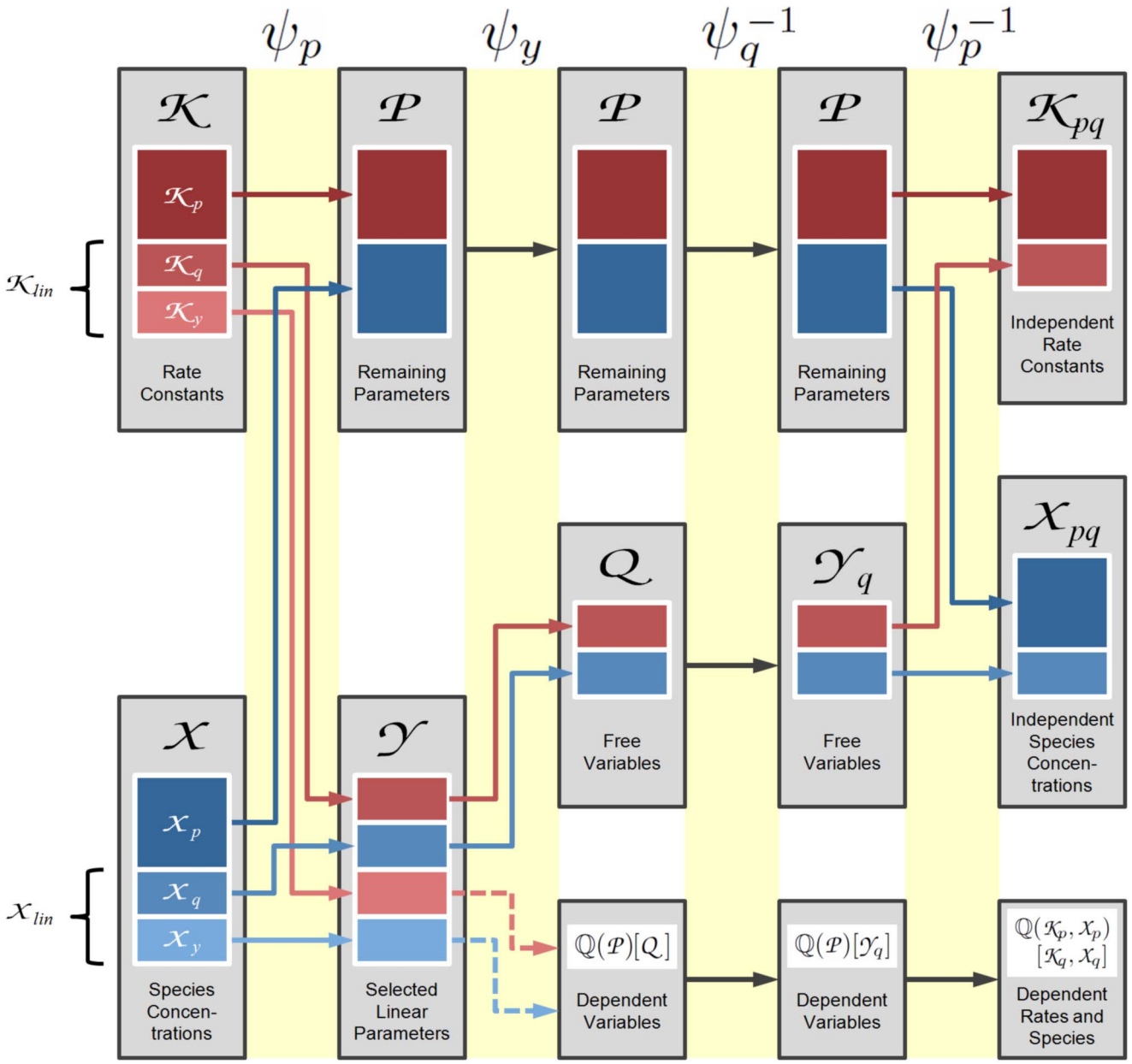


Figure 1. Overview of the *py*-substitution method. Quantities in a mass action model can be separated into kinetic rate constants (set $\mathcal{K}$, red) and species abundances or concentrations (set $\mathcal{X}$, blue). From $\mathcal{K} \cup \mathcal{X}$ a subset $\mathcal{K}_{lin} \cup \mathcal{X}_{lin}$ is selected on which all reaction velocities have only linear dependence. A function ψ_p maps these to elements in $\mathcal{Y}$ and the remaining $\mathcal{K}_p \cup \mathcal{X}_p$ to elements in $\mathcal{P}$. A second function ψ_y imposes the relations $\psi_p(\dot{\mathbf{x}}) = 0$ by expressing dependent variables in Y in terms of independent parameters $\mathcal{P} \cup \mathcal{Q}$. A third function, ψ_q^{-1}, is the inverse of ψ_y restricted to the independent parameters. The composition of ψ_p^{-1} with ψ_q^{-1} results in variables in $\mathcal{K}_y \cup \mathcal{X}_y$ being expressed in terms of variables in $\mathcal{K}_{pq} \cup \mathcal{X}_{pq}$, such that steady state is achieved. In the diagram, solid arrows are isomorphisms while dashed arrows are homomorphisms that replace dependent variables by equivalent expressions in independent parameters. See Table S1 for a complete listing of symbols and their meanings.
doi:10.1371/journal.pcbi.1002901.g001

If we extend ψ_{pq}^{-1} to $f \in \mathrm{span}_{\mathbb{Q}(K_p, X_p)}(\mathcal{K}_q, \mathcal{X}_q)$ homomorphically, we can compose ψ_{pq}^{-1} with ψ_{py},

$$\psi_{ss} = (\psi_{pq}^{-1} \circ \psi_{py}) : \begin{cases} \mathcal{K}_{pq} \to \mathcal{K}_{pq} \text{ (identity)} \\ \mathcal{X}_{pq} \to \mathcal{X}_{pq} \text{ (identity)} \\ \mathcal{K}_y \cup \mathcal{X}_y \to \mathrm{span}_{\mathbb{Q}(\mathcal{K}_p, \mathcal{X}_p)}(\mathcal{K}_q, \mathcal{X}_q) \end{cases}$$

The function ψ_{ss} then defines a map for which

$$\mathbf{S}\psi_{ss}(\mathbf{v}) = \mathbf{S}\bar{\mathbf{v}} = 0,$$

where steady state velocities in $\bar{\mathbf{v}}$ are in terms of elements in $\mathcal{K}$ and $\mathcal{X}$. A visual overview of the *py*-substitution method is given in Figure 1.

Results

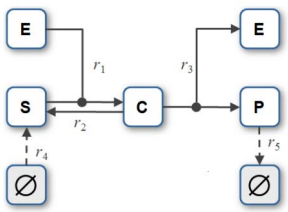
py-substitution permits flexible derivation of a steady state solution

An important goal in developing *py*-substitution was that it be generally applicable to any model whose reaction rates obey mass action kinetics. This requires that the independent quantities be chosen freely among the species concentrations and reaction rate constants, and that non-linear rate equations do not confound the derivation of a steady state expression. To demonstrate these capabilities we consider an open-system analog of the classical Michaelis-Menten model of enzyme kinetics (OMM, see also Figure 2). Substrate synthesis and product degradation allow this system to achieve a non-trivial steady state $\bar{\mathbf{v}} \neq \mathbf{0}$, which we derive here using four different substitution strategies. The set $\mathcal{R}$ of reactions for this model is given by

$$
\begin{aligned}
r_1 &: x_1 + x_2 \xrightarrow{v_1} x_3 \\
r_2 &: x_3 \xrightarrow{v_2} x_2 + x_1 \\
r_3 &: x_3 \xrightarrow{v_3} x_2 + x_4 \\
r_4 &: \emptyset \xrightarrow{v_4} x_1 \\
r_5 &: x_4 \xrightarrow{v_5} \emptyset
\end{aligned}
$$

The symbol $\emptyset$ represents a *source* or *sink* for mass and is not modeled by a time-varying species. From the set R we derive the stoichiometric matrix and reaction velocity vector,

$$
\mathbf{S} = \begin{bmatrix} -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & 1 & 0 & 0 \\ 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} k_1 x_1 x_2 \\ k_2 x_3 \\ k_3 x_3 \\ k_4 \\ k_5 x_4 \end{bmatrix}.
$$



Var	Symbol	Description
x_1	E	Enzyme
x_2	S	Substrate
x_3	C	Complex
x_4	P	Product

Figure 2. An open system analog of the classical Michaelis-Menten model for enzyme catalysis. Enzyme and substrate bind to form an intermediate complex, followed by catalysis and dissociation of the product. The substrate is synthesized by a zero-order reaction, r_4, and the product is degraded by a first-order reaction, r_5. See "omm1.m" in Protocol S1 for a complete description of the model.
doi:10.1371/journal.pcbi.1002901.g002

By Equation 4 this results in the following system of equations,

$$
\begin{aligned}
dx_1/dt &= -k_1 x_1 x_2 + k_2 x_3 + k_4 \\
dx_2/dt &= -k_1 x_1 x_2 + k_2 x_3 + k_3 x_3 \\
dx_3/dt &= k_1 x_1 x_2 - k_2 x_3 - k_3 x_3 \\
dx_4/dt &= k_3 x_3 - k_5 x_4
\end{aligned}
$$

for which we now derive functions ψ_{ss} such that $\mathbf{S}\psi_{ss}(\mathbf{v}) = \mathbf{S}\bar{\mathbf{v}} = 0$.

Homogeneous substitution: steady state concentrations do not uniquely determine reaction rate constants. The most straightforward substitution strategy is to let $\mathcal{K}_{lin} = \mathcal{K}$ and $\mathcal{X}_p = \mathcal{X}$. The corresponding function ψ_p maps

$$
\begin{aligned}
k_1 &\mapsto y_1 & x_1 &\mapsto p_1 \\
k_2 &\mapsto y_2 & x_2 &\mapsto p_2 \\
k_3 &\mapsto y_3 & x_3 &\mapsto p_3 \\
k_4 &\mapsto y_4 & x_4 &\mapsto p_4 \\
k_5 &\mapsto y_5
\end{aligned}
$$

See "omm1.m.trace.pdf" in Protocol S1 for details of this partition and all subsequent steps. Applying ψ_p to $\mathbf{v}$ results in a reaction velocity vector that is linear in $\mathbf{y}$, as required by Equation 17,

$$
\psi_p(\mathbf{v}) = \mathbf{P}\mathbf{y} = \begin{bmatrix} p_1 p_2 & 0 & 0 & 0 & 0 \\ 0 & p_3 & 0 & 0 & 0 \\ 0 & 0 & p_3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & p_4 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}.
$$

The resulting coefficient matrix is given by

$$
\mathbf{C} = \mathbf{S}\mathbf{P} = \begin{bmatrix} -p_1 p_2 & p_3 & p_3 & 0 & 0 \\ -p_1 p_2 & p_3 & 0 & 1 & 0 \\ p_1 p_2 & -p_3 & -p_3 & 0 & 0 \\ 0 & 0 & p_3 & 0 & -p_4 \end{bmatrix},
$$

which row reduces to

$$
\mathbf{C} \sim \mathbf{C}_{\text{rref}} = \begin{bmatrix} 1 & -p_3/(p_1 p_2) & 0 & 0 & -p_4/(p_1 p_2) \\ 0 & 0 & 1 & 0 & -p_4/p_3 \\ 0 & 0 & 0 & 1 & -p_4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (26)
$$

From Equation 26, we observe that $\text{rank}\ \mathbf{C} = 3$. Thus, of the 9 degrees of freedom in this system (5 rate constants plus 4 species concentrations), 3 will have values that are constrained by Equation 5. Since our substitution strategy only identifies 4 independent parameters, 2 additional elements mapped to $\mathcal{Y}$ must in fact be independent as well. These elements can be identified by the columns in $\mathbf{C}_{\text{rref}}$ that do not contain pivots, namely columns 2 and 5. To see this, note that Equation 26 yields the following basis for the null space of $\mathbf{C}$,

$$\mathbf{N}=\begin{bmatrix} p_3/(p_1p_2) & p_4/(p_1p_2) \\ 1 & 0 \\ 0 & p_4/p_3 \\ 0 & p_4 \\ 0 & 1 \end{bmatrix}.$$

Letting $\mathbf{q}=[q_1,q_2]^{\mathrm{T}}$, Equation 22 gives

$$\overline{\mathbf{y}}=\begin{bmatrix} (q_1p_3+q_2p_4)/(p_1p_2) \\ q_1 \\ (q_2p_4)/p_3 \\ q_2p_4 \\ q_2 \end{bmatrix}. \quad (27)$$

Thus by Equation 23, we have that $\mathcal{Y}_q=\{y_2,y_5\}$ and $\mathcal{Y}_q^c=\{y_1,y_3,y_4\}$. By Equation 24, Equation 27 can be described by the mapping function ψ_y :

$$\begin{aligned} p_1&\mapsto p_1 & y_1&\mapsto (q_1p_3+q_2p_4)/(p_1p_2) \\ p_2&\mapsto p_2 & y_2&\mapsto q_1 \\ p_3&\mapsto p_3 & y_3&\mapsto (q_2p_4)/p_3 \\ p_4&\mapsto p_4 & y_4&\mapsto q_2p_4 \\ & & y_5&\mapsto q_2 \end{aligned}$$

From ψ_y and ψ_p we construct the composite forward map, $\psi_{py}=(\psi_y\circ\psi_p)$:

$$\begin{aligned} k_1&\mapsto (q_1p_3+q_2p_4)/(p_1p_2) & x_1&\mapsto p_1 \\ k_2&\mapsto q_1 & x_2&\mapsto p_2 \\ k_3&\mapsto (q_2p_4)/p_3 & x_3&\mapsto p_3 \\ k_4&\mapsto q_2p_4 & x_4&\mapsto p_4 \\ k_5&\mapsto q_2 \end{aligned}$$

To reverse the substitution, notice from Equation 27 that $y_2=q_1$ and $y_5=q_2$, giving the following map, ψ_q^{-1}:

$$\begin{aligned} \mathcal{P}&\to\mathcal{P}(\text{identity}) & q_1&\mapsto y_2 \\ & & q_2&\mapsto y_5 \end{aligned}$$

This yields a composite backward map, $\psi_{qp}^{-1}=(\psi_p^{-1}\circ\psi_q^{-1})$:

$$\begin{aligned} p_1&\to x_1 & q_1&\mapsto k_2 \\ p_2&\to x_2 & q_2&\mapsto k_5 \\ p_3&\to x_3 & & \\ p_4&\to x_4 & & \end{aligned}$$

The complete steady state mapping $\psi_{ss}=(\psi_{qp}^{-1}\circ\psi_{py})$ is therefore

$$\begin{aligned} k_1&\mapsto (k_2x_3+k_5x_4)/(x_1x_2) & x_1&\mapsto x_1 \\ k_2&\mapsto k_2 & x_2&\mapsto x_2 \\ k_3&\mapsto k_5x_4/x_3 & x_3&\mapsto x_3 \\ k_4&\mapsto k_5x_4 & x_4&\mapsto x_4 \\ k_5&\mapsto k_5 & & \end{aligned} \quad (28)$$

Applying this transformation to the original vector of reaction velocities yields

$$\overline{\mathbf{v}}=\psi_{ss}(\mathbf{v})=\begin{bmatrix} k_2x_3+k_5x_4 \\ k_2x_3 \\ k_5x_4 \\ k_5x_4 \\ k_5x_4 \end{bmatrix},$$

which one can verify satisfies Equation 5. An interesting implication of this trivial application of *py*-substitution is that, because ψ_{py} maps every species concentration to an independent parameter, we can interpret Equation 28 to mean that any vector of steady state concentrations will be consistent with an infinite number of reaction rate constants. In this particular case, knowing all four concentrations tells us nothing about the rates of enzyme-substrate dissociation or product degradation. As we shall see, by using different substitution strategies, we have some flexibility in choosing which rate constants are constrained by the steady state concentrations, but the structure of the OMM model makes finding a unique set of rate constants impossible. In general, a unique set of reaction rate constants requires that the coefficient matrix be full rank, or

$$\text{rank } \mathbf{C}=d_y. \quad (29)$$

Since complete knowledge of the species concentrations implies $d_p=d_x$ and $d_y=d_k$, by Equation 20, Equation 29 becomes

$$\text{rank } \mathbf{S}=d_k.$$

In other words, a unique set of rate constants requires that the stoichiometric matrix be full rank, which is equivalent to requiring that the corresponding reaction network have no cycles. Since even a single reversible reaction represents a cycle, we conclude that in the general case, a set of steady state species concentrations does not imply a unique set of reaction rate constants.

Heterogeneous substitution: the number of independent model parameters is constant. Often, models contain species whose concentrations are difficult to measure or reactions whose rates have been well characterized. For such models it is preferable to partition sets $\mathcal{K}$ and $\mathcal{X}$ so that species whose concentrations are difficult to measure are mapped to $\mathcal{Y}$ while well-characterized reaction rates are mapped to $\mathcal{P}$. For example, if the kinetics of the enzyme are well characterized, an attractive partitioning of the OMM model might be $\mathcal{K}_p=\{k_1,k_2,k_3,k_5\}$ and $\mathcal{X}_p=\{x_2\}$. This yields a map ψ_p :

$$\begin{aligned} k_1 &\mapsto p_2 & x_1 &\mapsto y_1 \\ k_2 &\mapsto p_3 & x_2 &\mapsto p_1 \\ k_3 &\mapsto p_4 & x_3 &\mapsto y_2 \\ k_4 &\mapsto y_4 & x_4 &\mapsto y_3 \\ k_5 &\mapsto p_5 \end{aligned}$$

Again, see "omm2.m.trace.pdf" in Protocol S1 for complete details. Notice here that we have forced the enzyme kinetic parameters k_1, k_2, and k_3 to be independent by mapping them to elements in $\mathcal{P}$. The resulting coefficient matrix and null space basis are

$$\mathbf{C} = \begin{bmatrix} -p_1p_2 & p_3+p_4 & 0 & 0 \\ -p_1p_2 & p_3 & 0 & 1 \\ p_1p_2 & -(p_3+p_4) & 0 & 0 \\ 0 & p_4 & -p_5 & 0 \end{bmatrix},$$

$$\mathbf{N} = \begin{bmatrix} (p_3+p_4)/(p_1p_2p_4) \\ 1/p_4 \\ 1/p_5 \\ 1 \end{bmatrix},$$

which yield the steady state map ψ_{ss} :

$$\begin{aligned} k_1 &\mapsto k_1 & x_1 &\mapsto k_4(k_2+k_3)/(k_1k_3x_2) \\ k_2 &\mapsto k_2 & x_2 &\mapsto x_2 \\ k_3 &\mapsto k_3 & x_3 &\mapsto k_4/k_3 \\ k_4 &\mapsto k_4 & x_4 &\mapsto k_4/k_5 \\ k_5 &\mapsto k_5 \end{aligned}$$

As desired, x_2 is the only independent species concentration. Applying this transformation to the original vector of reaction velocities gives

$$\overline{\mathbf{v}} = \psi_{ss}(\mathbf{v}) = \begin{bmatrix} k_4(k_2+k_3)/k_3 \\ k_2k_4/k_3 \\ k_4 \\ k_4 \\ k_4 \end{bmatrix}.$$

Notice that even though the cardinality of $\mathcal{P}$ differs in this example as compared to the one above (5 *versus* 4), the cardinality of $\mathcal{Y}_q^c$ does not (3). Let d_c denote this cardinality. Obviously, $d_c = d_y - d_q$, or equivalently,

$$d_c = \text{ncols } \mathbf{C} - \text{ncols } \mathbf{N}.$$

This is simply the rank-nullity theorem again. By Equation 20, we can therefore conclude that

$$d_c = \text{rank } \mathbf{S}.$$

In other words, the final number of dependent elements in the steady state expression for a system is independent of the substitution strategy, and only depends on the structure of the reaction network.

Substitution with sublinear velocities: using *py*-substitution to resolve non-linearities (I). Some reaction velocities are zero-order but well-characterized. For example, if the rate v_4 of substrate synthesis in the OMM model has been accurately measured, we may wish to partition $\mathcal{K}$ such that $k_4 \in \mathcal{K}_p$. The resulting mapping function ψ_p, however, fails to linearize $\mathbf{v}$. To compensate, we introduce a pseudospecies $\hat{x}_5 = 1$ and let $v_4 = k_4\hat{x}_5$. If we now partition $\mathcal{X}$ such that $\hat{x}_5 \in \mathcal{X}_{lin}$, the linearity of $\psi_p(\mathbf{v})$ in $\mathbf{y}$ is preserved and we may continue as before.

To illustrate this approach, we again let $\mathcal{X}_p = \{x_1, x_2, x_3, x_4\}$ and $\mathcal{K}_p = \{k_4\}$. The remaining rate constants and one pseudospecies are partitioned into sets $\mathcal{K}_{lin}$ and $\mathcal{X}_{lin}$, respectively, such that $\psi_p(\hat{x}_5) = y_5$. See "omm3.m.trace.pdf" in Protocol S1 for details. The resulting velocity vector is linear and yields a coefficient matrix whose null space is two-dimensional,

$$\mathbf{N} = \begin{bmatrix} p_3/(p_1p_2) & p_5/(p_1p_2) \\ 1 & 0 \\ 0 & p_5/p_3 \\ 0 & p_5/p_4 \\ 0 & 1 \end{bmatrix}. \quad (30)$$

However, one of these two dimensions is constrained by the pseudospecies. We are thus not at liberty to take a general linear combination as per Equation 22 but must find $\mathbf{q}$ such that

$$\psi_{py}(\hat{x}_5) = 1. \quad (31)$$

By our choice of ψ_p, and by Equations 22 and 30, we have $\psi_{py}(\hat{x}_5) = (\mathbf{q})_2$. Equation 31 is therefore satisfied when $(\mathbf{q})_2 = 1$. This gives $\mathbf{q} = [q_1, 1]^{\mathrm{T}}$ and

$$\overline{\mathbf{y}} = \begin{bmatrix} (p_3q_1+p_5)/(p_1p_2) \\ q_1 \\ p_5/p_3 \\ p_5/p_4 \\ 1 \end{bmatrix}.$$

The complete steady state mapping $\psi_{ss} = (\psi_{qp}^{-1} \circ \psi_{py})$ is

$$\begin{aligned} k_1 &\mapsto (k_4+k_2x_3)/(x_1x_2) & x_1 &\mapsto x_1 \\ k_2 &\mapsto k_2 & x_2 &\mapsto x_2 \\ k_3 &\mapsto k_4/x_3 & x_3 &\mapsto x_3 \\ k_4 &\mapsto k_4 & x_4 &\mapsto x_4 \\ k_5 &\mapsto k_4/x_4 & \hat{x}_5 &\mapsto 1 \end{aligned}$$

As desired, k_4 remains an independent parameter. Applying this transformation to the original vector of reaction velocities yields

$$\overline{\mathbf{v}} = \psi_{ss}(\mathbf{v}) = \begin{bmatrix} k_4 + k_2 x_3 \\ k_2 x_3 \\ k_4 \\ k_4 \\ k_4 \end{bmatrix}.$$

Substitution with superlinear velocities: using *py*-substitution to resolve non-linearities (II). Some reaction velocities are superlinear in their reactant concentrations. If good estimates for these concentrations do not exist, we would like to partition these species into $\mathcal{X}_{lin}$. Analogous to the sublinear case above, doing so results in a velocity vector $\psi_p(\mathbf{v})$ that is non-linear in $\mathbf{y}$. Fortunately, the strategy above is useful here as well: introduce a pseudospecies for each superlinearity, calculate a basis for the null space of the coefficient matrix, and identify basis vector coefficients that satisfy the constraints imposed by the pseudospecies.

Let us consider a version of the OMM model where the rate of product formation is proportional to the square of the enzyme-substrate complex, $v_3 = k_3 x_3^2$. Let us further assume that no estimate exists for the value of x_3. We would therefore like $\psi_p(x_3) \in \mathcal{Y}$. Since this fails to linearize the velocity, we introduce a pseudospecies $\hat{x}_5 = x_3^2$ and let $v_3 = k_3 \hat{x}_5$. We now define ψ_p such that

$$\begin{array}{ll} k_1 \mapsto p_2 & x_1 \mapsto y_4 \\ k_2 \mapsto p_3 & x_2 \mapsto p_1 \\ k_3 \mapsto p_4 & x_3 \mapsto y_1 \\ k_4 \mapsto y_5 & x_4 \mapsto y_3 \\ k_5 \mapsto p_5 & \hat{x}_5 \mapsto y_2 \end{array}$$

This satisfies the linearity requirement and maps x_3 and $\hat{x}_5$ to the lowest indices in $\mathcal{Y}$, thereby favoring these quantities to become dependent parameters. See "omm4.m.trace.pdf" in Protocol S1 for details. The resulting coefficient matrix has a null space that is spanned by the columns of

$$\mathbf{N} = \begin{bmatrix} p_1 p_2 / p_3 & -1/p_3 \\ 0 & 1/p_4 \\ 0 & 1/p_5 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Letting $\mathbf{q} = [q_1, q_2]^T$ maps k_4 and x_1 to $\mathcal{Q}$ and satisfies our requirement that $\psi_{py}(x_3)$ and $\psi_{py}(\hat{x}_5) \in \mathcal{Y}_q^c$. As in the previous section, however, one dimension of $\mathbf{N}$ is constrained by the pseudospecies. Specifically, we require that $\psi_{py}(\hat{x}_5) = \psi_{py}(x_3^2)$. by Equation 22, this requires that

$$(q_2 - p_1 p_2 q_1)^2 / p_3^2 = q_2 / p_4.$$

Solving for q_1 (we may just as easily have solved for q_2; in this example, whether k_4 or x_1 map to $\mathcal{Q}$ is immaterial), we are left with the following:

$$\overline{\mathbf{y}} = \begin{bmatrix} \sqrt{q_2/p_4} \\ q_2/p_4 \\ q_2/p_5 \\ q_2(p_3\sqrt{q_2})/(p_1 p_2 \sqrt{p_4}) \\ q_2 \end{bmatrix}.$$

The complete steady state mapping $\psi_{ss} = (\psi_{qp}^{-1} \circ \psi_{py})$ is

$$\begin{array}{ll} k_1 \mapsto k_1 & x_1 \mapsto (k_4 + k_2\sqrt{k_4/k_3})/(k_1 x_2) \\ k_2 \mapsto k_2 & x_2 \mapsto x_2 \\ k_3 \mapsto k_3 & x_3 \mapsto \sqrt{k_4/k_3} \\ k_4 \mapsto k_4 & x_4 \mapsto k_4/k_5 \\ k_5 \mapsto k_5 & \hat{x}_5 \mapsto k_4/k_3 \end{array}$$

Applying this transformation to the original vector of reaction velocities yields

$$\overline{\mathbf{v}} = \psi_{ss}(\mathbf{v}) = \begin{bmatrix} k_4 + k_2\sqrt{k_4/k_3} \\ k_2\sqrt{k_4/k_3} \\ k_4 \\ k_4 \\ k_4 \end{bmatrix}.$$

This example illustrates that, using pseudospecies, a mapping function ψ_p can always be found such that ψ_y can be derived using linear methods. Non-linearities introduced by pseudospecies can then be resolved on a case-by-case basis, resulting in the final steady-state solution.

Py-substitution is more general, but not less efficient, than King-Altman

Some chemical reaction systems are linear in the species concentration vector, or can be rendered linear by assuming that the concentrations of certain species don't change over time. The classical model for malate synthesis is an example of the latter [42]. Here, the enzyme fumarase binds reversibly to fumarate and hydrogen in either order, followed by reversible binding of hydroxyl and reversible formation of malate (Figure 3). The reactions for this model are

$$\begin{array}{ll} r_1 : x_1 + x_6 \xrightarrow{v_1} x_3 & r_7 : x_3 \xrightarrow{v_7} x_1 + x_6 \\ r_2 : x_3 + x_7 \xrightarrow{v_2} x_4 & r_8 : x_4 \xrightarrow{v_8} x_3 + x_7 \\ r_3 : x_1 + x_7 \xrightarrow{v_3} x_5 & r_9 : x_5 \xrightarrow{v_9} x_1 + x_7 \\ r_4 : x_5 + x_6 \xrightarrow{v_4} x_4 & r_{10} : x_4 \xrightarrow{v_{10}} x_5 + x_6 \\ r_5 : x_4 + x_8 \xrightarrow{v_5} x_2 & r_{11} : x_2 \xrightarrow{v_{11}} x_4 + x_8 \\ r_6 : x_2 \xrightarrow{v_6} x_1 + x_9 & r_{12} : x_1 + x_9 \xrightarrow{v_{12}} x_2 \end{array}$$

By Equation 1, the corresponding reaction velocities are

$$\begin{aligned} v_1 &= k_1 x_1 x_6 & v_7 &= k_7 x_3 \\ v_2 &= k_2 x_3 x_7 & v_8 &= k_8 x_4 \\ v_3 &= k_3 x_1 x_7 & v_9 &= k_9 x_5 \\ v_4 &= k_4 x_5 x_6 & v_{10} &= k_{10} x_4 \\ v_5 &= k_5 x_4 x_8 & v_{11} &= k_{11} x_2 \\ v_6 &= k_6 x_2 & v_{12} &= k_{12} x_1 x_9 \end{aligned} \tag{32}$$

The stoichiometric matrix is

$$\mathbf{S} = \begin{bmatrix} -1 & 0 & -1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & -1 & 0 & 0 & -1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Notice that the submatrix formed by the first five rows of $\mathbf{S}$ satisfies the definition for a linear model given in Equation 7. Call this submatrix $\mathbf{S}_5$ and let $\mathbf{x}' = [x_1, \ldots, x_5]^{\mathrm{T}}$ be the vector of enzyme concentrations. If we assume that the substrate concentrations $x_6, \ldots, x_9$ are time-invariant, the steady state equation for this model becomes

$$\mathbf{S}_5 \mathbf{v} = 0. \tag{33}$$

Because $\mathbf{S}_5$ satisfies Equation 7, we may define the following transition rate constants

$$\begin{aligned} k'_{1,3} &= k_1 x_6 & k'_{3,1} &= k_7 \\ k'_{3,4} &= k_2 x_7 & k'_{4,3} &= k_8 \\ k'_{1,5} &= k_3 x_7 & k'_{5,1} &= k_9 \\ k'_{5,4} &= k_4 x_6 & k'_{3,4} &= k_{10} \\ k'_{4,2} &= k_5 x_8 & k'_{2,4} &= k_{11} \\ k'_{2,1} &= k_6 & k'_{1,2} &= k_{12} x_9 \end{aligned} \tag{34}$$

Substituting Equations 34 into 32 results in a velocity vector $\mathbf{v}'$ that is linear in $\mathbf{x}'$. Let $\mathbf{P}_{x'} = (\partial v'_i / \partial x'_j)$ as before, where $i = 1, \ldots, d_k$ and $j = 1, \ldots, d_{x'}$. This gives

$$\mathbf{v}' = \mathbf{P}_{x'} \mathbf{x}'. \tag{35}$$

If we now define a matrix

$$\mathbf{K} = \mathbf{S}_5 \mathbf{P}_{x'}, \tag{36}$$

Equation 33 becomes

$$\mathbf{K} \mathbf{x}' = 0, \tag{37}$$

where the elements of $\mathbf{K}$ are given in Equation 11. The solution to Equation 37 is given by Equation 13, which we saw may be evaluated using the King-Altman method. Alternatively, we may solve Equation 33 directly using *py*-substitution. Given that *py*-substitution applies to a more general class of mass action models then King-Altman, we wondered whether this flexibility came at the cost of computational efficiency. Here we show that, for models that can be treated using the King-Altman method, *py*-substitution yields an equivalent result, and at no loss of efficiency.

***Py*-substitution and King-Altman yield equivalent steady state expressions.** Equation 37 has been solved previously using *KAPattern* [26]. The solution is reproduced here in "fum1.m.trace.pdf" in Protocol S1. For each enzyme i, $1 \le i \le 5$, the steady state concentration has the form

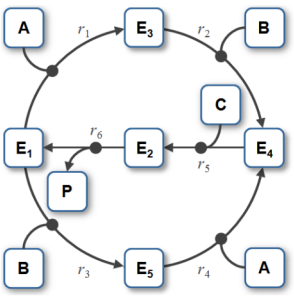
$$\bar{x}_i^{ka} = \frac{N_i^{ka}}{D^{ka}}. \tag{38}$$

In this subsection only, we use $\bar{x}_i^{ka}$ to mean the i^{th} element of the vector $\bar{\mathbf{x}}$, made to satisfy Equation 37 by the King-Altman method. The element $\bar{x}_i^{py}$ is defined analogously for *py*-substitution. To solve Equation 33 by *py*-substitution, we partition X into subsets

$$\begin{aligned} \mathcal{E} &= \{x_1, \ldots, x_5\} \\ \mathcal{E}^c &= \{x_6, \ldots, x_9\}, \end{aligned} \tag{39}$$

and define ψ_p such that

$$\psi_p : \begin{cases} \mathcal{K} \cup \mathcal{E}^c \to \mathcal{P} \\ \mathcal{E} \to \mathcal{Y}. \end{cases}$$



Var	Symbol	Description
x_1	E_1	Fumarase
x_2	E_2	E_1+A+B+C
x_3	E_3	E_1+A
x_4	E_4	E_1+A+B
x_5	E_5	E_1+B
x_6	A	Fumarate
x_7	B	Hydrogen
x_8	C	Hyrdoxyl
x_9	P	Malate

Figure 3. The model of malate synthesis used to compare *py*-substitution with the King-Altman method. This mechanism for the conversion of fumarate to malate by the enzyme fumarase was proposed in [42]. Fumarase binds to fumarate and hydrogen in either order, then hydroxyl, followed by formation of the product, malate. All reactions are reversible. See "fum1.m" in Protocol S1 for a complete description of the model.
doi:10.1371/journal.pcbi.1002901.g003

The resulting coefficient matrix is precisely the matrix of rate constants, **K**. The null space of **K** is one-dimensional and spanned by a single basis vector **n**. In our solution, the basis vector is normalized to element $(\mathbf{n})_5$, which by Equations 22 and 23 yield a partition of $\mathcal{Y}$ into subsets $\mathcal{Y}_q=\{y_5\}$ and $\mathcal{Y}_q^c=\{y_1,\ldots,y_4\}$. After reversing the substitution we find that the steady state concentration of each enzyme likewise has the form

$$\bar{x}_i^{py}=\frac{N_i^{py}}{D^{py}}. \tag{40}$$

By inspection, Equations 38 and 40 are related by the following:

$$N_i^{py}=\bar{x}_5^{py}N_i^{ka},\quad i=1,2,3,4 \tag{41}$$

$$D^{py}=N_5^{ka}. \tag{42}$$

In other words, the solutions given by *KAPattern* and *py*-substitution are not identical. The disparity arises from Equation 13, which imposes the constraint $\sum_i \bar{x}_i^{ka}=1$. When derived by King-Altman, the steady state expression for each enzyme is therefore a ratio of the total enzyme concentration. In contrast, *py*-substitution results in $\bar{x}_1^{py},\ldots,\bar{x}_4^{py}$ being expressed in terms of $\bar{x}_5^{py}$, the only element $x\in\mathcal{E}$ for which $\psi_{py}(x)\in\mathcal{Q}$. Despite this disparity, Equations 40 to 42 can be combined to give

$$\bar{x}_i^{py}=\bar{x}_5^{py}N_i^{ka}/N_5^{ka}.$$

Therefore,

$$\begin{aligned}\sum_{i=1}^{5}\bar{x}_i^{py}&=\frac{\bar{x}_5^{py}}{N_5^{ka}}\sum_{i=1}^{5}N_i^{ka}\\&=\bar{x}_5^{py}(D^{ka})/N_5^{ka}\\&=\bar{x}_5^{py}/\bar{x}_5^{ka}.\end{aligned}$$

If we likewise impose the constraint $\sum_i \bar{x}_i^{py}=1$, then $\bar{x}_5^{ka}=\bar{x}_5^{py}$, and for $i\neq 5$,

$$\begin{aligned}\bar{x}_i^{py}&=\bar{x}_5^{py}N_i^{py}/N_5^{ka}\\&=(N_5^{ka}/D^{ka})N_i^{ka}/N_5^{ka}\\&=N_i^{ka}/D^{ka}\\&=\bar{x}_i^{ka}.\end{aligned}$$

The two solutions are thus equivalent.

***Py*-substitution is not less efficient than King-Altman.** We next wondered whether the King-Altman method is computationally more efficient than direct algebraic solution of the linear steady state equation (Equation 18). The King-Altman method requires exhaustive enumeration of valid King-Altman patterns. The number of patterns depends critically on the structure of the model. A model of strongly connected species generates $d_x^{d_x-2}$ patterns while a simple cycle generates only d_x [38]. By comparison, solving Equation 18 requires Gaussian elimination on the matrix **K**. For a fixed-precision numeric matrix, this would take at most $O(d_x^3)$; however, since **K** has symbolic entries rather than numerical ones, the sizes of the entries grow with the number of row operations. In fact, as Equation 15 shows, the number of valid King-Altman patterns is precisely the number of terms in the polynomial expansion of the minors. Thus even a few species, if highly connected, can generate thousands of terms and easily overwhelm conventional memory architectures.

To evaluate the performance of *py*-substitution versus *KAPattern*, we generated random models with six species and anywhere from 10 to 20 first-order reactions between them. Three distinct realizations were generated for each model. Models for which *KAPattern* failed – typically because the stoichiometric matrix described a disjoint network – were discarded. The command-line version of Matlab 2010b was used to derive the steady state concentration vector for each model using *py*-substitution and *KAPattern*, and for *py*-substitution the command-line version of Maple 14 was used as well. Internal memory was cleared prior to each derivation to prevent caching. The architecture used was a commodity netbook PC running Windows XP SP3 with an Intel 1.7 GHz Atom processor and 1 GB RAM. The derivation was repeated in triplicate for each realization to reduce variance introduced by the CPU scheduler. Execution times include initialization of the symbolic variables and coefficient matrix, kernel calculation, and derivation of $\bar{\mathbf{y}}$ in the case of *py*-substitution, and all steps prior to file writing in the case of *KAPattern*.

Results from the simulation are given in Figure 4. The data show that using Matlab, *KAPattern* provides consistently better performance and better scaling with respect to the number of reactions. This is likely because *KAPattern* uses Wang algebra to avoid explicit representation of the fully expanded minors in memory [25]. In contrast, Gaussian elimination of the coefficient matrix uses MuPAD, the Matlab symbolic engine, which is memory intensive and sensitive to expression swell. Models of even modest degree exhaust physical memory and cause "thrashing", resulting in poor runtime performance for models larger than 15 reactions. However, using Maple, direct solution of the steady state equation is typically an order of magnitude faster than *KAPattern* and exhibits identical scaling. This is likely because Maple's symbolic solver considers equations in increasing order of their memory footprint. This data therefore argues that the King-Altman method is not more efficient than direct solution of the steady state equation.

***Py*-substitution is more general than King-Altman.** To solve the steady state equation, the King-Altman method requires that the stoichiometric matrix **S** satisfies Equation 7. As we saw above, for **S** to satisfy Equation 7 we must be able to partition $\mathcal{X}$ into two disjoint sets, a set $\mathcal{E}$ of "enzymes" and a complementary set $\mathcal{E}^c$ of "substrates". The partition must be such that every reaction $r\in\mathcal{R}$ consumes a single species in $\mathcal{E}$ and produces a single, different species in $\mathcal{E}$. All other species produced or consumed by r must be in $\mathcal{E}^c$. The concentrations of these substrates are assumed to be time-invariant. As such, rows in **S** that correspond to substrates can be removed, and the substrate concentrations can be incorporated into the kinetics of the reactions. By inspection, the only such partition for the fumarase model is Equation 39, analyzed above.

By comparison, *py*-substitution does not require that the stoichiometric matrix satisfies Equation 7. The substrates $x_6,\ldots,x_9$ can therefore remain variable with respect to time and incorporated into the steady state solution, of which there are many. Without recourse to pseudospecies, the six bimolecular reaction velocities require that x_6,x_7,x_9 and x_1,x_3,x_5 be partitioned separately into sets $\mathcal{X}_p$ and $\mathcal{X}_{lin}$, or vice-versa. One such partition is

$$\mathcal{K}_p = \{k_1, \ldots, k_5, k_{12}\}$$
$$\mathcal{K}_{lin} = \{k_6, \ldots, k_{11}\}$$
$$\mathcal{X}_p = \{x_1, \ldots, x_5\}$$
$$\mathcal{X}_{lin} = \{x_6, \ldots, x_9\}.$$

The resulting coefficient matrix has a five-dimensional null space, consistent with Equation 21 since $d_y = 10$ and $\mathrm{rank}\ \mathbf{S} = 5$. A basis for this null space is given by the columns in $\mathbf{N}$, where

$$\mathbf{N} = \begin{bmatrix} 0 & 0 & 0 & 0 & (p_6 p_7)/p_8 \\ -p_{10}/p_9 & (p_4 p_{11} + p_1 p_7)/p_9 & 0 & 0 & 0 \\ -1 & (p_4 p_{11})/p_{10} & (p_2 p_9)/p_{10} & 0 & 0 \\ p_{10}/p_{11} & -p_4 & (p_3 p_7)/p_{11} & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (p_5 p_{10})/p_8 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The set $\mathcal{Y}_q$ must therefore contain five elements. We may select these elements with some flexibility by our choice of basis vector coefficients. The simplest choice, $\mathbf{q} = [q_1, \ldots, q_5]^{\mathrm{T}}$, yields the steady state mapping ψ_{ss} :

$k_1 \mapsto k_1$ $\quad k_6 \mapsto (k_{12} x_1 x_9)/x_2$ $\quad x_1 \mapsto x_1$ $\quad x_6 \mapsto x_6$
$k_2 \mapsto k_2$ $\quad k_7 \mapsto (x_6(k_1 x_1 + k_4 x_5))/x_3 - (k_{10} x_4)/x_3$ $\quad x_2 \mapsto x_2$ $\quad x_7 \mapsto x_7$
$k_3 \mapsto k_3$ $\quad k_8 \mapsto (k_2 x_3 x_7 + k_4 x_5 x_6)/x_4 - k_{10}$ $\quad x_3 \mapsto x_3$ $\quad x_8 \mapsto x_8$
$k_4 \mapsto k_4$ $\quad k_9 \mapsto (k_{10} x_4 + k_3 x_1 x_7)/x_5 - k_4 x_6$ $\quad x_4 \mapsto x_4$ $\quad x_9 \mapsto x_9$
$k_5 \mapsto k_5$ $\quad k_{10} \mapsto k_{10}$ $\quad x_5 \mapsto x_5$
$k_{12} \mapsto k_{12}$ $\quad k_{11} \mapsto k_5 x_4 x_8 / x_2$

Other maps are available, however. By Equation 25, the submatrix formed by taking any 5 linearly independent rows of $\mathbf{N}$ produces a different vector of coefficients, and thus a different partition of $\mathcal{Y}$. For our particular choice of ψ_p above, 72 partitions are possible, calculated by testing which combinations of 5 rows in $\mathbf{N}$ are linearly independent. As an illustration, consider the case where the rate constants k_8 and k_{11} are easier to measure than substrates x_7 and x_8. Because of this, we would prefer x_7 and x_8 to be dependent variables. Equivalently, we want $\psi_{py}(x_7)$, $\psi_{py}(x_8) \in \mathcal{Y}_q^c$, and $\psi_{py}(k_8)$, $\psi_{py}(k_{11}) \in \mathcal{Y}_q$. Since $\psi_p(k_8) = y_3$ and $\psi_p(k_{11}) = y_6$, any 5×5 submatrix of $\mathbf{N}$ containing rows 3 and 6 whose determinant is not zero will accomplish this. Below is the vector $\mathbf{q}'$ calculated from the matrix formed by rows 3, 5, 6, 7, and 10.

$$\mathbf{q}' = \begin{bmatrix} q_2 \\ q_4 \\ (q_1 p_{10} + q_2 p_{10} - q_4 p_4 p_{11})/(p_2 p_9) \\ (q_3 p_8)/(p_5 p_{10}) \\ q_5 \end{bmatrix},$$

This results in the desired steady state mapping, ψ'_{ss} :ψ'_{ss} :

$k_1 \mapsto k_1$ $\quad x_1 \mapsto x_1$
$k_2 \mapsto k_2$ $\quad x_2 \mapsto x_2$
$k_3 \mapsto k_3$ $\quad x_3 \mapsto x_3$
$k_4 \mapsto k_4$ $\quad x_4 \mapsto x_4$
$k_5 \mapsto k_5$ $\quad x_5 \mapsto x_5$
$k_6 \mapsto (k_{12} x_1 x_9)/x_2$ $\quad x_6 \mapsto x_6$
$k_7 \mapsto (x_6(k_1 x_1 + k_4 x_5))/x_3 - (k_{10} x_4)/x_3$ $\quad x_7 \mapsto (k_{10} x_4 + k_8 x_4 - k_4 x_5 x_6)/(k_2 x_3)$
$k_8 \mapsto k_8$ $\quad x_8 \mapsto (k_{11} x_2)/(k_5 x_4)$
$k_9 \mapsto (k_{10} x_4)/x_5 - k_4 x_6 + (k_3 x_1 (k_{10} x_4 + k_8 x_4 - k_4 x_5 x_6))/(k_2 x_3 x_5)$ $\quad x_9 \mapsto x_9$
$k_{10} \mapsto k_{10}$
$k_{11} \mapsto k_{11}$
$k_{12} \mapsto k_{12}$

This offers another illustration of how the choice of substitution strategy and null space basis vectors allow one to choose independent parameters flexibly among sets $\mathcal{K}$ and $\mathcal{X}$ when solving for steady state. See "fum2.m.trace.pdf" in Protocol S1 for details of this derivation.

Steady state establishes a threshold for drug-induced cell death

Finally, we sought to use *py*-substitution to characterize the relationship between steady state and the response to the cancer drug, dulanermin. Dulanermin is a recombinant human form of the endogenous ligand TRAIL, whose mechanism for triggering cell death is modeled in version 1.0 of the *extrinsic apoptosis reaction model*, or EARM [14]. This model considers the biochemical events following engagement of the death receptors 4 and 5 (DR4/5), including receptor-induced cleavage of initiator caspases, positive-feedback by effector caspases, and feed-forward amplification by the mitochondrial pathway following outer membrane permeabilization, or MOMP (Figure 5). The EARM model was trained on data derived from HeLa cells co-treated with cyclohexamide, an inhibitor of protein synthesis that results in hypersensitivity to TRAIL [43]. Accordingly, any amount of ligand in the EARM model results in cell death. The abundance of ligand still affects the time of death, defined for example by the time t_{PARP} at which half of the caspase 3 target protein PARP has been cleaved (Figure 6A, left) [44]. Note in this section we refer to the abundance of a species rather than its concentration, as these are the units chosen by the original authors.

In the absence of cyclohexamide, however, HeLa cells do not all die following exposure to TRAIL. Rather, a fraction of cells persist, and this resistance is a function of the proteomic state prior to stimulation [6]. To capture this phenomenon, we extended the EARM model so that proteins continued to be synthesized and degraded following exposure to TRAIL. Specifically, we introduced 43 new synthesis and degradation fluxes as well as 2 protein inactivation reactions (see "xearm.mpl" in Protocol S1). These reactions were chosen so that every species is subject to at least one efflux. We refer to our extended model as xEARM. Because xEARM satisfies our definition of a mass action model, we use *py*-substitution to identify an analytical expression for its steady state. To derive this expression, a mapping function ψ_p was chosen so that every non-zero parameter in EARM was mapped to an independent parameter in P. As a result, we were able to preserve the snap-action dynamics of MOMP that is central to the original model (Figure 6A, right). Honoring the published parameters

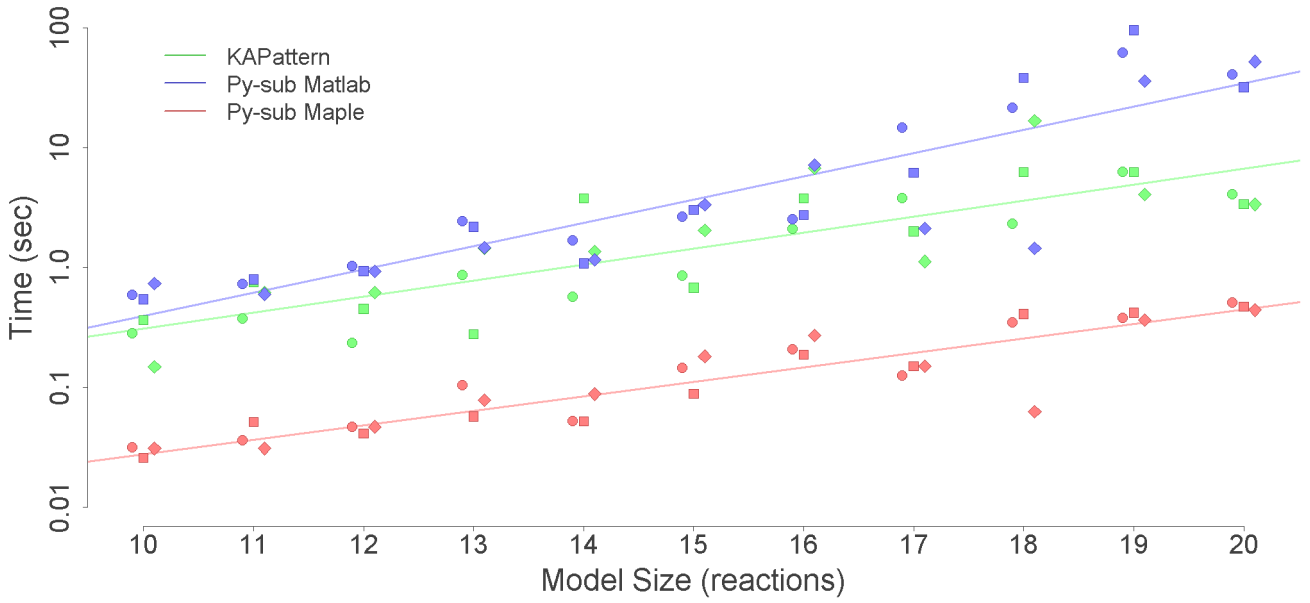


Figure 4. Computational performance of *KAPattern* versus *py*-substitution, implemented in either Matlab or Maple. Given a first-order model with six species and the number of reactions indicated by the x-axis, the time required to derive an expression for the steady state of the model is indicated by the y-axis. Three random realizations were used for every model size. Every calculation was performed in triplicate, but the error in calculation time was negligible.
doi:10.1371/journal.pcbi.1002901.g004

required that we introduce two pseudospecies, one for each the di- and tetrameric forms of Bax (variables x_{32} and x_{35}, respectively),

$$\hat{x}_{59} = x_{32}^2 \tag{43}$$

$$\hat{x}_{60} = x_{35}^2. \tag{44}$$

The coefficient matrix $\mathbf{C}$ and null space basis matrix $\mathbf{N}$ were calculated as before, with the latter calculation requiring less than a minute on our benchmark PC. The null space of $\mathbf{C}$ has 17-dimensions, resulting in a matrix of basis vectors of the form

$$\mathbf{N} = [\,\mathbf{n}_1 \quad \mathbf{n}_2 \quad \mathbf{n}_3 \quad \mathbf{n}_4 \quad \ldots \quad \mathbf{n}_{17}\,].$$

Basis vectors $\mathbf{n}_4$ to $\mathbf{n}_{17}$ preserve the steady state ratios of paired synthesis and degradation reactions. Vector $\mathbf{n}_{17}$, for example, ensures that a change δ in k_{114} results in a change δx_{42} in k_{113}, where x_{42} is the abundance of Cytochrome C in the mitochondria and k_{113} and k_{114} are its rates of synthesis and degradation, respectively. The vector $\mathbf{n}_3$ scales the steady state abundances of mitochondrial Bax and Bcl2 complexes with respect to changes in the rate of Bcl2 synthesis. Vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ are algebraically intractable and thus defy simple biochemical interpretation. Two of these vectors, $\mathbf{n}_1$ and $\mathbf{n}_3$, are constrained by the pseudospecies $\hat{x}_{59}$ and $\hat{x}_{60}$. To resolve these constraints, note that Equations 43 and 44 require that

$$\psi_p(\hat{x}_{59}) = \psi_p(x_{32}^2) \tag{45}$$

$$\psi_p(\hat{x}_{60}) = \psi_p(x_{35}^2). \tag{46}$$

By our mapping function ψ_{py} (see "xearm.mpl.trace.pdf" in Protocol S1, pp. 120–121), Equations 45 and 46 become

$$y_{42} = y_{21}^2 \tag{47}$$

$$y_{43} = y_{23}^2, \tag{48}$$

where $\psi_y(y_{21}), \psi_y(y_{23}), \psi_y(y_{42})$, and $\psi_y(y_{43}) \in \mathrm{span}_{\mathbb{Q}(\mathcal{P})}(\{q_1, q_2, q_3\})$. Solving Equation 48 for q_3 gives

$$q_3 = \frac{b_1}{b_2} q_2^2, \tag{49}$$

where $b_1, b_2 \in \mathbb{Q}[\mathcal{P}]$. Substituting Equation 49 into Equation 47 and solving for q_1 gives

$$q_1 = \frac{-a_2 \pm \sqrt{a_2^2 - 4a_1a_3}}{2a_3}, \tag{50}$$

where a_1, a_2, and $a_3 \in \mathbb{Q}(\mathcal{P})[q_2]$ (see "xearm.mpl.trace.pdf" in Protocol S1, pp. 121–126).

Obviously, Equation 50 identifies an explicit bistability in the xEARM model. Basis vector coefficient q_1 — and by Equation 49, q_3 — can take either of two values for any numerical realization of the model. By examination of ψ_{py}, we find that these two coefficients affect all modified and compound species, as well as synthesis rates for proteins within and upstream of the mitochondria. Using the parameter values supplied in [14], however, we

find that one of the solutions to Equation 50 is negative. The corresponding steady state is therefore infeasible and the solution was discarded.

In addition to parameters in [14], a full numerical realization of the xEARM model requires values for parameters $p_{71},\ldots,p_{86}$ and $q_2,q_4,\ldots,q_{17}$. All but three of these elements represent first-order degradation rate constants, to which we assigned values equivalent to a half-life of one hour. This value was based on global quantifications of protein turnover in mammalian cells, which revealed that signaling proteins tend to be short-lived [45]. Two of the elements, p_{77} and p_{78}, represent first-order inactivation fluxes, which we assumed to be ten times faster than protein degradation. The final element q_2 is the steady state abundance of the mitochondrial Bax2:Bcl2 complex, which we set to 20 molecules. Six of the elements were then modified from their initial values to better match the dynamics of caspase activation and PARP cleavage, as reported in [14]. The complete table of parameter values required to initialize and numerically integrate the xEARM model is given in Table S2.

For comparison, Table S3 lists the steady state abundances of species in the original and extended EARM models, sorted in order of decreasing difference. As expected, every species in EARM with a non-zero abundance has precisely the same abundance in xEARM, since these are independent parameters in the steady state solution. Among species with zero abundance in EARM, the mitochondrial Bax:Bcl2 complex exhibits the greatest disparity, with the steady state abundance in xEARM being in the low thousands of molecules. Ubiquitinated, cleaved caspase 3 and cleaved PARP are also in the low hundreds of molecules, but this represents only a small fraction of their total cellular abundance. A full 25 species with zero abundance in the EARM model have an abundance of less than 1 molecule in xEARM. This indicates that, even though the steady state reaction velocities are markedly different between EARM and xEARM, by using *py*-substitution we were able to engineer a steady state where the species abundances are appreciably similar between the two models.

Next we asked whether the xEARM model remained viable in the presence of low doses of TRAIL, but still exhibited MOMP when stimulated with high doses of TRAIL. To do so we created a numerical realization of the model using the parameters from Table S2, then perturbed the model from its steady state using a step increase in the abundance of TRAIL (variable x_1). The magnitude of the step ranged from 1 to 100-fold and was followed by numerical integration of the mass balance equations out to 48 hours. As shown in Figure 6A, MOMP is only observed in xEARM when TRAIL is increased by $10^{1.25}$-fold or more. We label this minimum dose of TRAIL required for MOMP L_{thresh}. Increments less than L_{thresh} result in a small and transient change in cleaved PARP abundance, followed by a return to the pre-stimulated steady state. By comparison, any magnitude dose of TRAIL causes MOMP in the original EARM model.

This ability of xEARM to distinguish between low and high doses of TRAIL, in conjunction with an analytical expression for its steady state, allowed us to systematically perturb the steady state and ask how these perturbations affect the sensitivity to TRAIL. To illustrate this capability we varied the steady state abundance of each major xEARM species over a 100-fold range, centered about each species' wildtype value as reported in Table S2. For each variation, we performed a binary search to identify L_{thresh}. The results from this procedure are plotted in Figure 6B. As expected, increases in XIAP, Bcl2, FLIP, and Bar result in reduced sensitivity to TRAIL stimulation, while increases in Procaspase 8, TRAIL receptor DR4/5, Bax, and Bid result in increased sensitivity [46]. What is interesting, however, is the following. First, TRAIL sensitivity is most affected by changes in the abundance of Procaspase 8 and Bar, an inhibitor of active caspase 8 [47]. The ability to activate caspase 8, then, appears to be a critical determinant of TRAIL sensitivity, as previously suggested [48,49]. Second, the abundances of Procaspase 3, 6, and 9 have little effect on the sensitivity to TRAIL. This observation is in good agreement with the model-based prediction that induction of MOMP does not require positive-feedback via this caspase loop [14].

A common metric for describing how model parameters affect the sensitivity to TRAIL is to calculate the change in time at which death occurs in response to a small change in each parameter [6,44,50]. It is conceivable, however, that changes in the time of death do not accurately reflect changes in the threshold of TRAIL at which death occurs. Therefore, to test this assumption we calculated parameter sensitivity coefficients for the ligand threshold, $\partial L_{\text{thresh}}/\partial p$, and the time at which death occurs, $\partial t_{\text{PARP}}/\partial p$, using the xEARM and EARM models, respectively. The numerators $\partial L_{\text{thresh}}$ and ∂t_{PARP} were calculated by backward finite difference approximation and all sensitivities were normalized to the maximum observed sensitivity for each metric (Figure 6C). The data show good agreement for positive regulators of TRAIL sensitivity, but some disparity in the negative regulators. Specifically, while t_{PARP} is particularly sensitive to changes in XIAP and Bcl2, L_{thresh} is most sensitive to changes in Bar. This result argues that some caution should be taken when equating changes in the time of death with changes in TRAIL sensitivity.

Discussion

We have described a simple but flexible method for deriving analytical expressions for the steady states of mass action models. Central to our method is the observation that mass action models are systems of polynomial equations that are generally no greater than degree 2. This permits a partitioning of rate constants and species concentrations into disjoint sets of quantities, $\mathcal{P}$ and $\mathcal{Y}$, where the reaction velocity vector is linear with respect to the variables in $\mathcal{Y}$. If the cardinality of $\mathcal{Y}$ is greater than the rank of the stoichiometric matrix, then the steady state equation can be solved analytically using simple linear methods.

There is considerable benefit to deriving an analytical expression for the steady state of a model. An analytical expression can be used to identify network ultrasensitivity [51], robustness [52], multistationarity [53], and invariants [54]. For enzyme catalytic models that have no true steady state but nevertheless satisfy the assumptions for quasi-steady state, an analytical expression can relate the rate of product formation to the initial concentrations of the substrates and enzyme [55]. Critically, these properties do not depend on the numerical values of the parameters, which may be difficult to measure [56]. In our companion manuscript, we show that analytical steady state expressions can be used to identify changes in the kinetic rate constants that do not alter the species concentrations. These *isostatic perturbations* can be used to characterize the dynamic plasticity of a system, and also how changes in the rates of protein turnover can affect the response to perturbation, independently of changes to steady state concentrations.

Even if numerical interrogation is ultimately intended and all parameters must be assigned values, deriving an analytical expression for the steady state still confers a number of benefits. First, including steady state constraints can facilitate the construction of a model [57]. As illustrated by our treatment of the Open Michaelis-Menten model, *py*-substitution affords considerable flexibility in selecting which quantities are independent — thus requiring numerical values prior to simulation — and which quantities can be derived from the independent quantities. This partly transforms the problem of

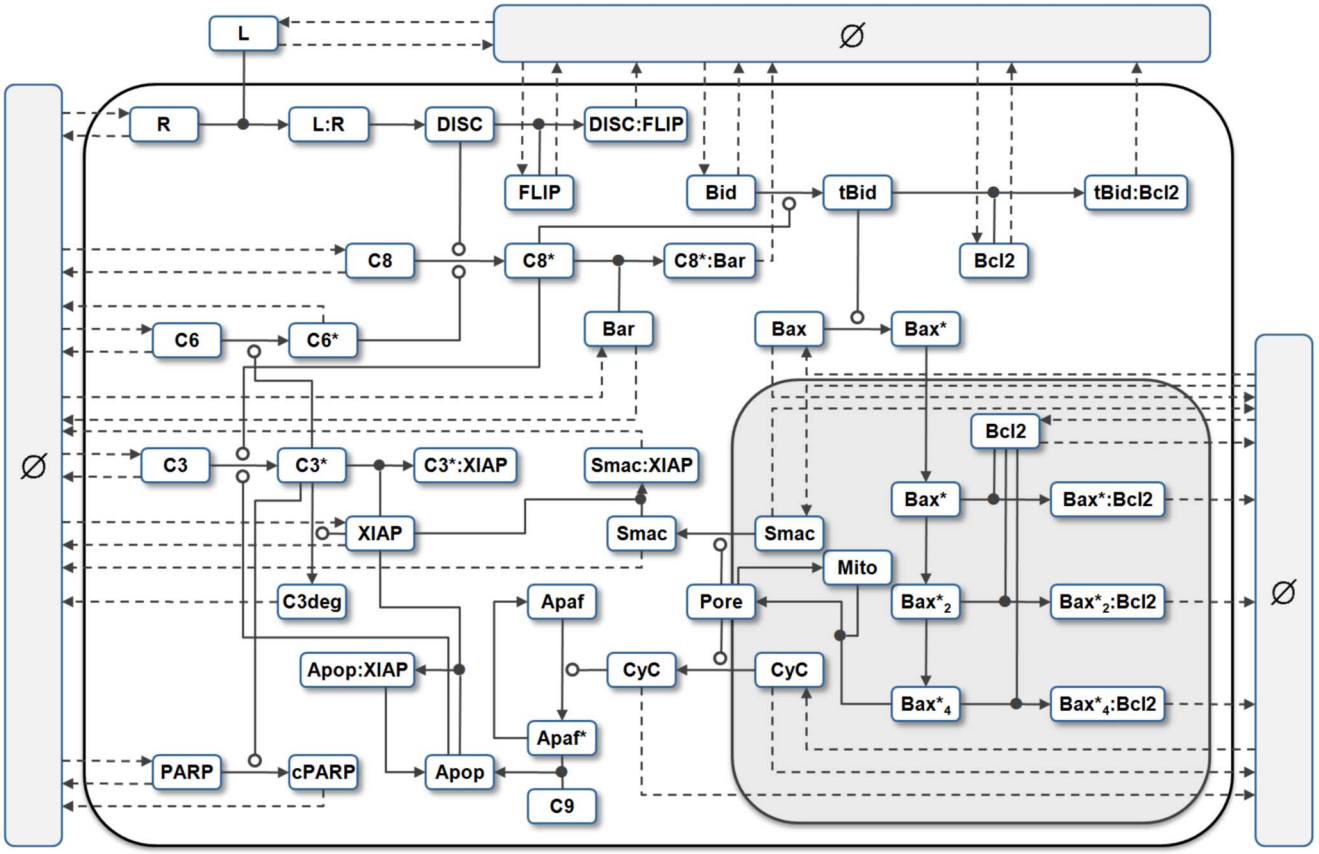


Figure 5. Reaction diagram for the xEARM model. Reactions new to this version include all fluxes to or from a source node, indicated by dashed lines to or from a Ø. In addition, the activation of *Apaf* was made reversible, as were the formation of mitochondrial pores. The complete model contains 58 species and 115 reactions. See [14] for a description of the original EARM model, and "xearm.mpl" in Protocol S1 for a complete description of xEARM.
doi:10.1371/journal.pcbi.1002901.g005

parameterizing a model from one of numerically fitting the rate constants to available data [58], to one of identifying the steady state expression that maximizes incorporation of known quantities into the independent set of parameters. Second, incorporating steady state concentration measurements can reduce the total number of parameters required. In the traditional approach to parameterization, every rate constant is assigned a value prior to simulation, as well as the abundance of any species not subject to synthesis and degradation. Using *py*-substitution, only independent quantities must be assigned a value. This number is equal to the total number of species and reactions, minus the rank of the stoichiometric matrix. As the stoichiometric matrix approaches full rank, this number converges to the number of species. Since most systems have more reactions than species, *py*-substitution often requires fewer parameters than the traditional approach. This can be observed in the xEARM model, where 119 parameters are required for simulation after deriving a steady state expression using *py*-substitution (100 rate constants, 18 species, and the mitochondrial volume), versus 133 parameters required for traditional parameterization (115 rate constants, 17 species, and the mitochondrial volume).

Further, in the case of the xEARM model, we have demonstrated that an analytical expression of the steady state allows systematic characterization of its effect on the response to perturbation. This was made possible in two ways. First, it allowed the model to operate at a non-trivial steady state. In the original EARM model, infinite sensitivity to TRAIL is caused by unbalanced reactions. Once the receptor is engaged, caspase cleavage and pore formation proceed deterministically to completion. As a result, for cells to be "alive" prior to stimulation, the model must assume a trivial steady state in which the abundance of TRAIL and all reaction velocities are zero. Using *py*-substitution, we were able to engineer a non-trivial steady state that is viable at low doses of TRAIL. Second, we were able to apply systematic changes to the steady state concentrations. By virtue of the mapping function ψ_{ss}, these resulted in compensating changes to the kinetic rate constants such that steady state was preserved. For each modification, we were then able to calculate the number of TRAIL molecules required to induce cell death, as well as the sensitivity of this threshold to changes in the steady state concentrations of different species.

Previous studies with models operating at trivial steady states employed sensitivity metrics that were with respect to the time at which death occurs, and not whether it occurs [6,44]. These studies suggested that the dynamics of TRAIL-induced cell death depend critically on Bcl-2 [44]. Also, whether cell death proceeds to completion depends on XIAP [44], and whether the mitochondrial feed-forward loop is required depends on the ratio of XIAP to Procaspase 3 [59]. In contrast, our analysis indicates that whether cell death occurs is primarily determined by the ratio of Procaspase 8 to its negative regulator, Bar. Our sensitivity analysis with respect to the threshold at which death occurs is therefore related to but distinct from analyses that consider only

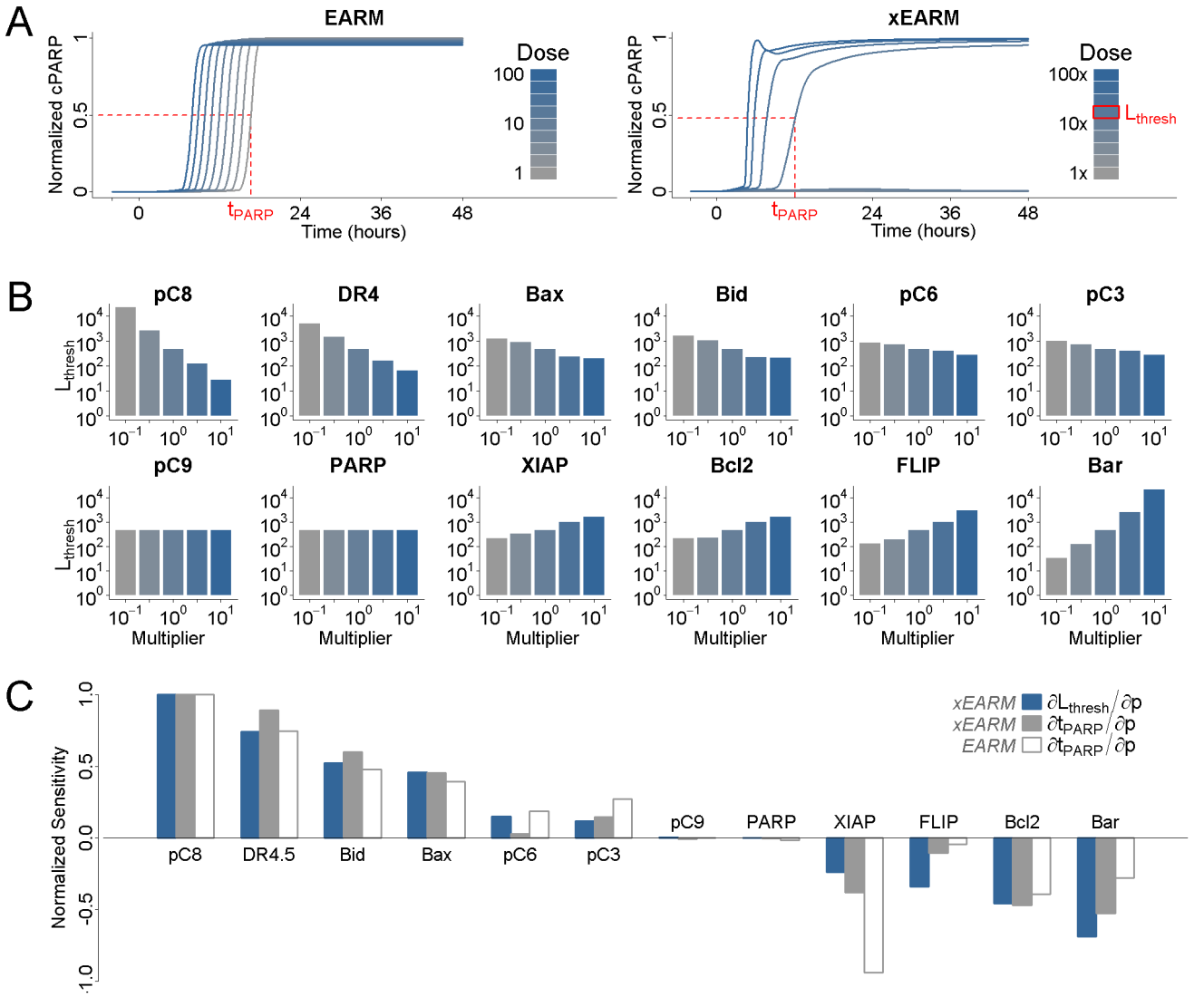


Figure 6. Determinants of sensitivity of TRAIL-induced cell death. (A) The dynamics of PARP cleavage are shown for EARM (left) and xEARM (right), in response to increasing doses of the TRAIL ligand (gray to blue). The abundance of cleaved PARP for each model has been normalized to the maximum observed abundance. For each model, for a particular dose of TRAIL, the time t_{PARP} at which PARP is 50% cleaved is indicated by the dashed red lines. For xEARM, the minimum abundance of TRAIL required to observe MOMP, L_{thresh}, is indicated on the color scale at right. (B) Changes in L_{thresh} in response to changes in the steady state abundance of 12 primary xEARM species. Species have been sorted from left to right in order of those for which an increase in abundance results in the greatest increase in TRAIL sensitivity, to those for which an increase in abundance result in the greatest decrease in sensitivity. (C) Normalized sensitivity coefficients for L_{thresh}, calculated using the xEARM model (blue), and t_{PARP}, calculated using both EARM (white) and xEARM (gray), for each of the 12 primary species in (B).
doi:10.1371/journal.pcbi.1002901.g006

the timing of death, and may relate better to clinical applications since we don't assume co-treatment with cyclohexamide.

For all these reasons, an analytical expression for the steady state of a model can be of general benefit to cell systems modeling. Indeed, other methods have previously addressed the challenge of deriving analytical steady state expressions, most notably the King-Altman method. Prior to the advent of modern computers, the authors realized that for a particular class of mass action models, the laborious calculation of steady state enzyme ratios could be achieved by a conceptually simpler graphical method. As we have shown, however, this simpler approach is no longer more efficient. More significantly, the King-Altman method requires that all reactions be first- or pseudo-first order in the time-varying species. Without this stipulation, Equation 7 no longer holds and the reaction network can no longer be described by a graph. This requirement is often stated as a pair of assumptions: 1) that no enzyme is itself a substrate and 2) that all substrates remain constant over the time scale of steady state formation [22]. The second of these can be considered common to any method that treats time-varying species as constants when solving the steady state equation. The first of these, however, is violated by any cascade of post-translational modifications, for example the well-known MAP kinase cascade [60].

Although recent methods relax these assumptions [28,29], in the contemporary systems biology literature, analytical derivation of the steady state rarely, if ever, precedes numerical interrogation of a model. Since this derivation is of considerable value, we sought to develop a method that was simple, scalable, and general to mass

action models. First, we described our method using only concepts from linear algebra, and we have provided complete code for all seven examples described in this manuscript, with implementations in either Matlab or Maple. Second, we show that *py*-substitution scales well. The xEARM model has 58 species and 115 reactions, and we were able to derive a steady state expression in less than a minute on a conventional desktop computer. Finally, we demonstrated that *py*-substitution can be generally applied to chemical reaction networks whose reaction velocities are modeled by mass action kinetics. This is a considerably broader class of models than can be addressed using the King-Altman and other methods, which require that the reaction network exhibit specific structural properties.

This does, however, open up an interesting avenue for further research: precisely what properties must a mass action model exhibit for its steady state to be derived using *py*-substitution ? How many different steady state expressions are possible, and which of these is the "best"? As we have shown with the fumarase model, even after the rate constants and species concentrations were partitioned into sets $\mathcal{P}$ and $\mathcal{Y}$, 72 different steady state expressions were possible. These different expressions arose from flexibility in selecting the pivot columns in the coefficient matrix, since the pivot vs. free columns partition the linear variables into dependent vs. independent variables. Equivalently, these different expressions arise from flexibility in ordering the linear variables, since different orderings permute the columns of the coefficient matrix and result in a different reduced row echelon form. Since the number of possible steady state expressions is large but finite, a combinatorial optimization strategy ought to be able to identify the best steady state expression, where the difference between any two expressions could take into account measurement uncertainty in the independent quantities, as well as computational complexity in deriving the final steady state expression.

Finally, we consider that the steady state may not be the only state of interest, but perhaps specified dynamic states as well. Essentially, this replaces the zero vector in Equation 5 with a vector of non-zero values. From linear algebra, we know that the solution to this dynamic equation can be expressed as the sum of a particular solution to the dynamic equation and an arbitrary point in the null space of the coefficient matrix. The solution is thus straightforward, raising the possibility of incorporating specific dynamic states into the parameterization of a model as well.

Supporting Information

Protocol S1 This zip file contains all source code required to run *py*-substitution, in either Matlab or Maple. Implementations for every model described in this manuscript, including a pdf trace of the steady state derivation using *py*-substitution, are also included. For further details please see the included README file.
(ZIP)

Table S1 This table provides a summary and description of all mathematical symbols used in this manuscript.
(PDF)

Table S2 This table compares the non-trivial steady state abundances of all molecular species in the xEARM model with their counterparts in the EARM model, published in [14]. All abundances are in molecules.
(PDF)

Table S3 This table gives values for all parameters required to numerically integrate the xEARM model. Also see "xearm.mpl" in Supporting Protocol S1.
(PDF)

Acknowledgments

The authors thank Ronan Fleming for insightful suggestions.

Author Contributions

Conceived and designed the experiments: PML GT AH. Performed the experiments: PML GT. Analyzed the data: PML GT. Wrote the paper: PML GT AH.

References

1. Werner SL, Barken D, Hoffmann A (2005) Stimulus specificity of gene expression programs determined by temporal control of IKK activity. Science 309: 1857–61.
2. Nakakuki T, Birtwistle MR, Saeki Y, Yumoto N, Ide K, et al. (2010) Ligand-specific c-Fos expression emerges from the spatiotemporal control of ErbB network dynamics. Cell 141: 884–96.
3. Batchelor E, Loewer A, Mock C, Lahav G (2011) Stimulus-dependent dynamics of p53 in single cells. Molecular Systems Biology 7: 488.
4. Miller-Jensen K, Janes Ka, Wong YL, Griffith LG, Lauffenburger Da (2006) Adenoviral vector saturates Akt pro-survival signaling and blocks insulin-mediated rescue of tumor necrosis-factorinduced apoptosis. Journal of Cell Science 119: 3788–98.
5. O'Dea EL, Kearns JD, Hoffmann A (2008) UV as an amplifier rather than inducer of NF-kappaB activity. Molecular Cell 30: 632–41.
6. Spencer SL, Gaudet S, Albeck JG, Burke JM, Sorger PK (2009) Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. Nature 459: 428–32.
7. Ashkenazi A, Herbst RS (2008) To kill a tumor cell: the potential of proapoptotic receptor agonists. The Journal of Clinical Investigation 118: 1979–90.
8. Mahalingam D, Oldenhuis CNaM, Szegezdi E, Giles FJ, de Vries EGE, et al. (2011) Targeting TRAIL towards the clinic. Current Drug Targets 12: 2079–90.
9. Niepel M, Spencer SL, Sorger PK (2009) Non-genetic cell-to-cell variability and the consequences for pharmacology. Current Opinion in Chemical Biology 13: 556–61.
10. Tyson JJ, Chen KC, Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Current Opinion in Cell Biology 15: 221–31.
11. Goldstein B, Faeder JR, Hlavacek WS (2004) Mathematical and computational models of immune-receptor signalling. Nature Reviews Immunology 4: 445–56.
12. Cheong R, Hoffmann A, Levchenko A (2008) Understanding NF-kappaB signaling via mathematical modeling. Molecular Systems Biology 4: 192.
13. Kearns JD, Hoffmann A (2009) Integrating computational and biochemical studies to explore mechanisms in NF-fkappagB signaling. The Journal of Biological Chemistry 284: 5439–43.
14. Albeck JG, Burke JM, Spencer SL, Lauffenburger DA, Sorger PK (2008) Modeling a snap-action, variable-delay switch controlling extrinsic cell death. PLoS Biology 6: 2831–52.
15. Monroe JG (2004) Ligand-independent tonic signaling in B-cell receptor function. Current Opinion in Immunology 16: 288–95.
16. Gough DJ, Messina NL, Clarke CJ, Johnstone RW, Levy DE (2012) Constitutive Type I Interferon Modulates Homeostatic Balance through Tonic Signaling. Immunity 36: 166–174.
17. Macia J, Regot S, Peeters T, Conde N, Solé R, et al. (2009) Dynamic signaling in the Hog1 MAPK pathway relies on high basal signal transduction. Science Signaling 2: ra13.
18. Artomov M, Kardar M, Chakraborty AK (2010) Only signaling modules that discriminate sharply between stimulatory and nonstimulatory inputs require basal signaling for fast cellular responses. The Journal of Chemical Physics 133: 105101.
19. Singh DK, Ku CJ, Wichaidit C, Steininger RJ, Wu LF, et al. (2010) Patterns of basal signalling heterogeneity can distinguish cellular populations with different drug sensitivities. Molecular Systems Biology 6: 369.
20. Hoffmann A, Levchenko A, Scott ML, Baltimore D (2002) The IkappaB-NF-kappaB signalling module: temporal control and selective gene activation. Science 298: 1241–5.
21. King E, Altman C (1956) A schematic method of deriving the rate laws for enzyme-catalyzed reactions. The Journal of Physical Chemistry 60: 1375–1378.
22. Thomson M, Gunawardena J (2009) The rational parameterization theorem for multisite posttranslational modification systems. Journal of Theoretical Biology 261: 626–36.
23. Volkenstein M, Goldstein B (1966) A new method for solving the problems of the stationary kinetics of enzymological reactions. Biochimica et Biophysica Acta 115: 471–477.
24. Cha S (1968) A simple method for derivation of rate equations for enzyme-catalyzed reactions under the rapid equilibrium assumption or combined assumptions of equilibrium and steady state. The Journal of Biological Chemistry 243: 820–5.

25. Lam CF, Priest DG (1972) Enzyme kinetics. Systematic generation of valid King-Altman patterns. Biophysical Journal 12: 248–56.
26. Qi F, Dash RK, Han Y, Beard Da (2009) Generating rate equations for complex enzyme systems by a computer-assisted systematic method. BMC Bioinformatics 10: 238.
27. Dickenstein A, Shiu A (2011) Chemical Reaction Systems with Toric Steady States. Bulletin of Mathematical Biology : 1–29.
28. Feliu E, Knudsen M, Andersen LN, Wiuf C (2012) An algebraic approach to signaling cascades with N layers. Bulletin of Mathematical Biology 74: 45–72.
29. Feliu E, Wiuf C (2012) Variable elimination in post-translational modification reaction networks with mass-action kinetics. Journal of Mathematical Biology 66: 281–310.
30. Sreenath SN, Cho KH, Wellstead P (2008) Modelling the dynamics of signalling pathways. Essays in Biochemistry 45: 1–28.
31. Palsson BO (2006) Systems Biology: Properties of Reconstructed Networks. Cambridge: Cambridge University Press. 334 pp.
32. Oberhardt MA, Chavali AK, Papin JA (2009) Flux balance analysis: interrogating genome-scale metabolic networks. Methods in Molecular Biology 500: 61–80.
33. Gianchandani EP, Chavali AK, Papin Ja (2010) The application of ux balance analysis in systems biology. Wiley Interdisciplinary Reviews Systems Biology And Medicine 2: 372–82.
34. Orth JD, Thiele I, Palsson BO (2010) What is flux balance analysis? Nature Biotechnology 28: 245–8.
35. Famili I, Mahadevan R, Palsson BO (2005) k-Cone analysis: determining all candidate values for kinetic parameters on a network scale. Biophysical Journal 88: 1616–25.
36. Jamshidi N, Palsson BO (2008) Formulating genome-scale kinetic models in the post-genome era. Molecular Systems Biology 4: 171.
37. Fromm SJ, Fromm HJ (1999) A two-step computer-assisted method for deriving steady-state rate equations. Biochemical and Biophysical Research Communications 265: 448–52.
38. Poland D (1989) King-Altman-Hill diagram method for open systems. The Journal of Physical Chemistry 93: 3605–3612.
39. Cox D, Little J, OShea D (2007) Ideals, Varieties, and Algorithms. Undergraduate Texts in Mathematics. Third edition. New York, NY: Springer New York. 551 pp. doi:10.1007/978-0-387-35651-8. URL http://www.springerlink.com/index/10.1007/978-0-387-35651-8.
40. Manrai AK, Gunawardena J (2008) The geometry of multisite phosphorylation. Biophysical journal 95: 5533–43.
41. Martínez-Forero I, Peláez-López A, Villoslada P (2010) Steady state detection of chemical reaction networks using a simplified analytical method. PloS One 5: e10823.
42. Hansen JN, Dinovo EC, Boyer PD (1969) Initial and equilibrium 18O, 14C, 3H, and 2H exchange rates as probes of the fumarase reaction mechanism. The Journal of Biological Chemistry 244: 6270–9.
43. Mühlenbeck F, Haas E, Schwenzer R, Schubert G, Grell M, et al. (1998) TRAIL/Apo2L activates c-Jun NH2-terminal kinase (JNK) via caspase-dependent and caspase-independent pathways. The Journal of Biological Chemistry 273: 33091–8.
44. Gaudet S, Spencer SL, Chen WW, Sorger PK (2012) Exploring the Contextual Sensitivity of Factors that Determine Cell-to-Cell Variability in Receptor-Mediated Apoptosis. PLoS Computational Biology 8: e1002482.
45. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, et al. (2011) Global quantification of mammalian gene expression control. Nature 473: 337–42.
46. Zhang L, Fang B (2005) Mechanisms of resistance to TRAIL-induced apoptosis in cancer. Cancer Gene Therapy 12: 228–37.
47. Zhang H, Xu Q, Krajewski S, Krajewska M, Xie Z, et al. (2000) BAR: An apoptosis regulator at the intersection of caspases and Bcl-2 family proteins. Proceedings of the National Academy of Sciences of the United States of America 97: 2597–602.
48. Eggert A, Grotzer MA, Zuzak TJ, Wiewrodt BR, Ho R, et al. (2001) Resistance to tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-induced apoptosis in neuroblastoma cells correlates with a loss of caspase-8 expression. Cancer Research 61: 1314–9.
49. Ganten TM, Haas TL, Sykora J, Stahl H, Sprick MR, et al. (2004) Enhanced caspase-8 recruitment to and activation at the DISC is critical for sensitisation of human hepatocellular carcinoma cells to TRAIL-induced apoptosis by chemotherapeutic drugs. Cell Death and Differentiation 11 Suppl 1: S86–96.
50. Schliemann M, Bullinger E, Borchers S, Allgower F, Findeisen R, et al. (2011) Heterogeneity Reduces Sensitivity of Cell Death for TNF-Stimuli. BMC Systems Biology 5: 204.
51. Knudsen M, Feliu E, Wiuf C (2012) Exact analysis of intrinsic qualitative features of phosphorelays using mathematical models. Journal of Theoretical Biology 300: 7–18.
52. Shinar G, Feinberg M (2010) Structural sources of robustness in biochemical reaction networks. Science 327: 1389–91.
53. Thomson M, Gunawardena J (2009) Unlimited multistability in multisite phosphorylation systems. Nature 460: 274–7.
54. Gunawardena J (2007) Distributivity and processivity in multisite phosphorylation can be distinguished through steady-state invariants. Biophysical Journal 93: 3828–34.
55. Segel LA (1988) On the validity of the steady state assumption of enzyme kinetics. Bulletin of mathematical biology 50: 579–93.
56. Bailey JE (2001) Complex biology with no parameters. Nature Biotechnology 19: 503–4.
57. Purvis JE, Radhakrishnan R, Diamond SL (2009) Steady-state kinetic modeling constrains cellular resting states and dynamic behavior. PLoS Computational Biology 5: e1000298.
58. Chen WW, Schoeberl B, Jasper PJ, Niepel M, Nielsen UB, et al. (2009) Input-output behaviour of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. Molecular Systems Biology 5: 239.
59. Aldridge BB, Gaudet S, Lauffenburger Da, Sorger PK (2011) Lyapunov exponents and phase diagrams reveal multi-factorial control over TRAIL-induced apoptosis. Molecular Systems Biology 7: 553.
60. Huang CY, Ferrell JE (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. Proceedings of the National Academy of Sciences of the United States of America 93: 10078–83.